

Compromise on arms control?

Recent noises from the United States and the Soviet Union suggest that the prospect of a deal on arms control may be more immediate than it has seemed.

MORE than usual confusion seems to mark the ending a week ago of the latest session of the bilateral talks on arms control at Geneva. On the face of things, President Reagan's declaration that the United States would not necessarily be bound by the SALT II agreement on strategic arms from the autumn of this year made the prospect of an agreement at Geneva less easily attainable than since the negotiations began. Yet Mr Mikhail Gorbachev, the general secretary of the Soviet Communist Party, appears to have sent Mr Reagan a conciliatory letter, coinciding with the summer break at Geneva, whose effect is to suggest that compromise is within the bounds of possibility on the control of weapons of intermediate range in Europe. What is going on? Were the Soviet Union's fierce reactions to Mr Reagan's announcement mere bluff? Or has the Soviet Union been told privately that it is Mr Reagan who was bluffing, and that his threat that SALT II might be abandoned was directed not so much at Moscow as at the hard-liners in Washington?

Enough has now leaked out about the conduct of the Geneva negotiations to sustain a degree of nervous optimism. Mr Gorbachev's message bears on one stumbling block: how to legislate for the complete banning of intermediate-range missiles (SS20 Soviet missiles and the Pershing II and cruise missiles deployed in Europe by the United States) while recognizing that SS20s could be quickly moved from Asia, that the British and French nuclear forces exist and that both sides also have a large number of shorter-range weapons which, if the major Soviet and US missiles were withdrawn, would acquire a strategic significance they now lack? Both sides at Geneva appear to have been captivated by the notion that it would be simplest to do away with all missiles of intermediate range (which, Mr Reagan's "zero option" of 1982, is progress in itself). Mr Gorbachev's suggestion now that compromise is possible suggests the possibility of an agreement on the basis of a sharply reduced but non-zero intermediate force. Few would complain at that.

Much less has been said about progress at Geneva on strategic arms. It would help, of course, if the two sides were merely to agree to limit strategic weapons to the numbers allowed under SALT II (a gigantic 5,000 warheads each) and were to take the opportunity to tidy up the treaty, making it acceptable to the US Congress, but that will not satisfy the aspirations of the rest of the world, while even the negotiators at Geneva appear to think that would be a prize not worth having (which, again, is progress of a kind). The best bet is that a fifty per cent reduction of warheads is attainable now; Mr Gorbachev made an important mark earlier this year with his declaration that even such an agreement could be only an interim measure, a prelude to further more substantial reductions of nuclear arms. The paradox is that, if that principle were accepted by both sides, the degree of reduction agreed in the next few months would be less important, and so would be more easily attained.

The third item on the agenda at Geneva is the most taxing, and that most likely to prevent an agreement being reached: to what extent will the devotion of the US administration to the Strategic Defense Initiative (SDI) stick in the throat of the Soviet Union? Here too, Mr Gorbachev has been making conciliatory noises: if research means only laboratory research, all

well and good. But then, Mr Gorbachev continues (as in his latest letter to Mr Reagan) why should not both sides reaffirm their adherence to the Anti-Ballistic Missile Treaty, undertaking that they will not quit the treaty for, say, 15 years? This position represents a Soviet retreat from the fierce declarations from Moscow in 1983, suggesting that merely toying with the idea of a defence against ballistic missiles was an act of aggression. But it seems improbable that Mr Gorbachev's moderation of the past few weeks suggest he will to let SDI go ahead full tilt.

Fortunately, even here constructive compromises should be possible. First, Mr Gorbachev probably knows as well as anybody of the reluctance of the US Congress to give the project the huge sums of money it has been asking for, implying that decisions about the deployment of SDI will not take place until long after the promised decision time in the early 1990s. Second, it is also well known that there has to be a general election in the United States two years from now, and that a new president (Mr Reagan cannot stand for a third term), whatever his or her political stripe, may not set as much store by the benefits of SDI. Finally, however meagre the outlook for this ambitious project, there is no obvious reason why the most likely outcome of this system, an efficient system of early warning based on infrared detectors, should not be welcomed by both sides. That is one sense in which Mr Gorbachev's suggestion that both sides should reaffirm the ABM treaty for 15 years is already a useful compromise. It requires from the United States only that degree of restraint likely to be forced on it by the technological difficulties of this ambitious project and by the increasing scepticism of the US Congress; if the two sides can make some kind of agreement this year, Congress will be even less willing to spend, with the result that the dangerous features of SDI will wither away.

So much for the carrots; Mr Gorbachev's stick is the threat that he will not attend a second summit meeting unless there is something substantial about which to agree. Once the summer is over, and the campaign for the mid-term elections has begun, there will be powerful forces in the United States pushing the administration towards some lasting achievement on the international front. To fail to have a summit meeting will seem, or be made to seem, a black mark in the first week in November. Not for the first time, Mr Reagan may find that his natural (and evidently deeply felt) suspicion of arms control agreements will have to be forsaken in the face of the general feeling that there is no more laudable goal. This is the turnabout forced on him during the first year of his administration six years ago. □

Research democracy

Moving Greenwich to Cambridge is not the end of the world, but users' voices need to be heard.

How should grant-making agencies manage their relations with those parts of the research community that rely on them for support? That is the question most urgently prompted by the decision of the British Science and Engineering Research Council (SERC) last week that the Royal Greenwich Observatory (RGO) should be moved to Cambridge rather than to



Edinburgh or Manchester.

The decision itself is impeccable, and properly within SERC's gift. SERC decided early in the year that a move was necessary, largely on the grounds that RGO had become a somewhat self-preoccupied institution that would benefit from stronger links with British universities. The managerial case for combining RGO with the Royal Observatory, Edinburgh, SERC's first preference, has been talked down by the British community of astronomers, while the more daring option, that of moving to Manchester, seems to have been abandoned out of deference to those among the RGO staff who believe that life north of the River Trent would be uncomfortable, a little like living in a foreign country. The result is that Cambridge will become both the administrative and academic centre of British astronomy. (Those now working at Edinburgh will increasingly think themselves marooned.)

This prospect is not necessarily bad; Cambridge is an excellent place. SERC also deserves some credit for having made a decision, a decision of any kind. But it has had to do so in the face of a torrent of complaints. Having pulled off this one triumph, SERC should resolve that it will not allow itself to be alienated a second time from those whose interests it exists to foster. This moral needs to be learned elsewhere as well.

Nobody doubts that SERC has the interests of British astronomers at heart. Over the past few years, it has been exceedingly generous towards them. There is now a prospect that British astronomers will have access to optical observing equipment, in the Canary Islands and in Hawaii, relatively more modern than they have enjoyed since, in the eighteenth century, Herschel made his name as a builder of telescopes. So why do British astronomers bite the hand that feeds them so well? Because they have a sense that SERC, acknowledged to be generous, is nevertheless acting high-handedly. It is not so much a benefactor as a nanny who "knows best". Why else does it not publish the two reports on the future of RGO (and Edinburgh) which apparently failed to come to an acceptable conclusion (in the eyes of SERC).

The remedy should be obvious. SERC should devise means by which its constituents can make their needs felt. Traditionally, SERC has delegated this task to the Space, Astronomy and Radio Board, formally taken as representative of astronomers because, in due course, most people of distinction pop up amongst its membership. But this is no longer an acceptable way of dealing with major decisions about the development of new instruments, telescopes in particular. Instead, there should be a committee whose members are not bound by secrecy and which is not forever having to measure its claims on resources by the calculation that rebuffs will diminish its authority in the competition for dwindling resources. Far better that there should be an open independent committee representing astronomers and their interests to which all working astronomers would have access, and able to ask for projects that are occasionally turned down. This is how high-energy physics is dealt with in the United States (see *Nature* 321, 636; 1986). The system does not rid high-energy physics of frustration, but it makes decisions intelligible.

How would such a system work in Britain, where restraints on public discussion seem all too natural? As things are the best lightning conductor for the opinions of the users of SERC's generously equipped new observatories would be a committee organized by the Royal Society or on some other independent base whose job would be the definition and repeated redefinition of researchers' needs. It would remain for SERC to decide how its facilities should be managed, which is its proper function. While about the job of setting up this mechanism, the research councils should also pay attention to the needs of other disgruntled parts of the research enterprise, the geophysicists, for example. The principle, which applies generally, is that users' needs are not merely useful planning information for managers, but are the only basis on which a sensible programme can be designed. □

Eureka discovered

The ministerial meeting of the Eureka project has said all the right things. Does it mean them?

THE European Eureka project, conceived by President François Mitterrand of France as a counter to what seemed the technological threat to Europe of the Strategic Defense Initiative, seems to have taken on a life of its own. The British government, for one, seems to have been converted from scepticism to enthusiasm, at least to judge from the speech at the opening of the third ministerial conference, in London this week, by the British Prime Minister, Mrs Margaret Thatcher. The British government seems to have been persuaded by the experience of the past eighteen months that the Eureka project has not flowered into yet another trans-national bureaucracy in Europe, that there is no danger that it will become a vehicle for other peoples' chauvinism and that it may even be a useful vehicle for urging on faltering Europe some of the policies from which it has consistently shied over recent years. That is more than a little to be grateful for.

Yet Eureka remains a project whose success so far are intangible, and whose promise is hard to pin down. Mrs Thatcher this week repeated the old adage that Britain is strong on discovery and weak in their application. She went on to complain at the tendency of technical people in Europe to disparage the application of good ideas, and to urge that people should pay more attention to the design of products suited to their intended markets. The difficulty with this familiar homily is that, now, it is out of date. Technical people all over Europe would give their eye teeth for the opportunity to design marketable products. They know that nations that fail on that score become impoverished remarkably quickly. The impediments to European prosperity are now more probably structural than psychological.

Structural? When much of Europe is linked together in a community called popularly a "common market"? How can that be? Because the common market is in no real way a common market in the ordinary sense, but rather a loose gathering of ten European nations provided with common (and often irksome) services by a central bureaucracy, with a general understanding that they may discriminate against each other's goods and services only by bureaucratic means and, to the extent that they are united, drawn together by their mutual suspicion of each other as traders. The unexpected success of Eureka, which functions by facilitating but not financing collaborative projects involving European companies and universities on technical projects, is that it has helped in several subtle ways to exorcise this long-standing suspicion.

So where will Eureka lead? Mrs Thatcher was right to insist this week that Eureka by itself is insufficient, and that something has to be done to make the common market genuinely common. But how? Governments' own xenophobia in public purchasing is an obvious place to start (as she said), but governments have been more willing to applaud the abstract principle than to change their own ways. That is why Eureka's well-wishers must also hope to build on the psychological changes now well under way in the informal European community, that in which people enjoy crossing frontiers without formality. One of the missing ingredients of Mrs Thatcher's speech this week was the recognition of the importance of the educational system, and especially of European universities, in providing the cultural cement that will, given encouragement, turn into economic cement as well. For, when everything has been said about the need that Europe should be more conscious of markets in Europe and elsewhere, it remains the case that Europe is dangerously short of the skills needed to carry out that task. Eureka may be working well so far, but its further success will only serve to draw attention to the urgent need for a coordinated system of higher education and research. This will be a great opportunity, but will at first be seen as a threat to national integrity. □

West German space research

Hermes causes sacking and cabinet unrest

Hamburg

A row over the West German space programme has resulted in the sacking of the head of the Department for Space Affairs, Aeronautics and Traffic Research. Wolfgang Finke, a Social Democrat, was removed from the post he has held for 13 years by the Minister of Research and Technology, Heinz Riesenhuber, a Christian Democrat, last week.

The two have disagreed over almost everything: Riesenhuber accused Finke of not having done enough for the ocean research programme; Finke in turn opposed his boss's decision to cut traffic research. But the crucial point of their disagreement reflects a political quarrel within the coalition cabinet over West German participation in European space programmes and, in particular, over participation in the French Hermes space shuttle project.

Like some ministers, Finke has fought for "stronger engagement for an independent European space potential that represents our political importance, our economic capacity, and technical knowledge".

Riesenhuber, however, is against a European space adventure. He has already agreed to cooperate in the construction of the Ariane 5 rocket, under French management, and to join with the United States to build the space station Columbus, at a cost for West Germany of DM 4 million. But two years ago he rejected France's offer of a share in the development of Hermes because of doubts over costs.

The French estimate of a total cost of DM 6,600 million was referred to as "total nonsense" by the minister, who considers that the worldwide satellite system necessary for the control and supervision of the space shuttle will alone cost at least DM 6,600 million. And he is very much aware of the words of Gerhard Stoltenberg, Minister of Finance, who announced to the cabinet in June last year that there was to be no money for Hermes or similar big projects. Before he is willing to reconsider his position, Riesenhuber wants a more realistic calculation of the likely costs.

On the other hand, Chancellor Helmut Kohl, another Christian Democrat, supports cooperation with France in general, and Foreign Minister Hans-Dietrich Genscher, of the Free Democrats, is urging his colleagues to back Hermes. The two — among others — have been influenced by a report published last week by the Research Institute for Foreign Policy, in which support is demanded for both

Hermes and a European space station. One of the authors of the paper is none other than fired Wolfgang Finke. It appears that Finke had told Kohl that he had his chief's support for the report — without the latter's knowledge. Riesenhuber thus had no choice but to sack him.

On balance there remains lukewarm support for the Hermes project within the cabinet. But questions over real costs and the extent of French domination of the project will need to be answered before any real commitment is forthcoming. **Jürgen Neffe**

Hermes limps on to ESA's agenda

WEST Germany gave a grudging commitment last week to go ahead with "preparatory studies" for a possible European space shuttle, Hermes, despite the sacking of the official responsible for coordinating Germany's space policy (see above).

However, the agreement last week by all 13 member states of the European Space Agency (ESA) that Hermes, so far only a French project, should be thus "Europeanized", falls far short of a total commitment to build a shuttle.

No money has yet been committed, but in three months, at the October meeting in Paris of the ESA council, states will be expected to make a declaration of what share they will take of a 48-million-ECU (European Currency Unit) nine-month detailed analysis of the real costs, benefits and likely technical problems of building Hermes.

ESA officials believe that the commitments made in October to this £32-million preparatory programme should be a "good indication" of likely future national involvement in the full construction of the space plane. However, a spokesman for the West German research ministry, BMFT, took time off from a European technology conference in London on Monday to say that German estimates of the true cost of Hermes were "five to six times" the £1,500 million or so suggested by the French. Moreover, the spokesman said, the German research minister, Heinz Riesenhuber, had told his British opposite number, Geoffrey Pattie, that Germany backed the British alternative project, "Hotol", a project for a horizontal take-off, partly air-breathing space-plane. "Hermes is old technology", the spokesman said. "Hotol is the future". **Robert Walgate**

Biotechnology

UNIDO centre taking shape

An international biotechnology centre for the developing countries came a step closer to realization last week with the appointment of its first director. Professor Irwin Gunsalus, emeritus professor at the University of Illinois, was chosen to head the International Centre for Genetic Engineering and Biotechnology (ICGEB) at a meeting organized by the United Nations Industrial Development Organization (UNIDO).

Five years have passed since the idea of a centre to bring the benefits of biotechnology to the developing countries was first proposed. During that time, many scientists, recognizing that there is a void in knowledge of molecular genetics and recombinant DNA technology in these countries, have given the centre their blessing. Now, UNIDO has provided funding to employ Gunsalus for three years, during which time he will recruit the scientists and technical staff needed to make the centre operational. He will also have to try to raise funds for the centre, which is not expected to be easy.

Italy has so far been generous in supporting the scheme and one of the centre's two branches is planned for Trieste. The other is to go to New Delhi. Rapid progress is unlikely, for almost all of the 37 countries that have signed up to support the centre are from the developing regions of Africa, South America, the Caribbean and Asia and are not easily able to provide money. The major industrial European countries, the United States and Japan have yet to support the centre and may well not do so until they have a chance of seeing how successful it will be.

Control of the centre is intended to remain within the developing countries so they can be confident that it is not just a route by which their best researchers can vanish to the advanced nations. Indeed, it is hoped that the flow will be in the opposite direction, with researchers from the advanced nations taking visiting fellowships at the centre. It is planned that each will have a staff of 30 scientists, 20 post-doctoral students and 40 technicians. Many of the staff will be on 2-3 year appointments.

Running the pair of institutes is not going to be easy. In theory, with the developing countries' heavy dependence on agriculture, biotechnology should bring enormous benefits. But much basic research on tropical agriculture remains to be done and the institutes will need to stimulate such research, at the same time fighting off demands for rapid returns on investment. **Alun Anderson**

US research

Campus bonanza from the military

Washington

US SECRETARY of Defense Caspar Weinberger last week announced the first round of winners in a \$110 million bonanza designed to improve universities' military research capabilities. The University Research Initiative (URI) will support 86 research projects at 70 institutions with awards ranging in value between \$170,000 and \$3 million.

The projects supported cover a wide spectrum. Most are intended to bring together small groups of researchers from different disciplines. But funds will also be set aside for instrumentation and for fellowships designed to encourage exchanges between universities and other institutions, especially the Department of Defense (DoD)'s own laboratories. Awards are conditional upon the successful conclusion of negotiations between DoD and the institutions.

The universities' enthusiastic response to URI seems to suggest that sensitivity on campus to military research has all but vanished. According to Dr Ronald Kerber, deputy under secretary of defence for research and advanced technology, most of the 963 URI research proposals received were of strikingly high quality. Kerber sees the large number of proposals received — worth nearly \$6,000 million in funds — as confirming universities' willingness to conduct basic research of possible military interest. Kerber said there had been no evidence of boycotts by academics opposed to the growth in the defence research budget.

Awards are typically in the \$1 million–\$2 million range, less than the \$4 million originally expected. That, Kerber explains, was partly to prevent faculty from "building empires". A single faculty member can be well supported with about \$200,000, so the average group will be of between five and ten researchers. Several awards will go to consortia of two or more universities.

Awards were made on the basis of competitive review by DoD scientists, helped by some outside consultants. As expected, most go to materials science and electronics.

Composite materials and electronic control dominate; there is also special emphasis on artificial intelligence, robotics, biotechnology, oceanography, hydrodynamics and biological structures. The awards, lasting 3–5 years, will be made through the Army Research Office, the Office of Naval Research, the Air Force Office of Scientific Research and the Defense Advanced Projects Research Agency.

The awards are still conditional upon availability of funds in DoD's budget for

fiscal year 1987, but Congress has so far strongly supported URI providing funds beyond those requested. The first URI round comes out of 1986 and 1987 funds; the next round will not be until 1988. All URI projects will be unclassified, and Kerber expects them to remain so.

Tim Beardsley

Scraping the barrel

Washington

IRONICALLY, on the day that US Secretary of Defense Caspar Weinberger announced the first competitive awards to be made under DoD's new University Research Initiative, the Senate finally gave in to pressure from the House of Representatives to divert \$55.6 million of DoD funds to unreviewed research and construction projects at nine named universities and colleges.

The funds are provided in amendments to an emergency DoD money bill. Ten such "pork-barrel" projects were approved at the end of last year. But as they had not been subjected to peer review, they were opposed by DoD itself, which prefers to make research awards by competitive merit review (see *Nature* 321, 549; 1986).

Several of the universities provided for in the DoD bill employed influential lobbying and public relations companies in Washington to make their case to wavering congressmen. Their task was made easier because federal support for university facilities is generally agreed to be inadequate, and politicians like to be seen appearing to benefit education in their home states.

Two attempts were made in the Senate to block the disputed projects, but the Senate finally voted 56–42 last Thursday to accept a House version of the bill that reinstated nine pork-barrel projects. The Senate Appropriations Committee's report notes, however, that the Senate does not intend to support similar unreviewed projects in future.

The original ten pork-barrel projects in the DoD appropriation bill were reduced to nine last December when Frank Rhodes, president of Cornell University, refused to accept \$10 million that Congress had earmarked for computer research there, unless the project was approved by merit review.

The universities due to benefit from Congress's generosity are: Wichita State (\$5 million); Nevada (\$3.5 million); Oklahoma State (\$1 million); Iowa State (\$6.5 million); Rochester Institute of Technology (\$11.1 million); Syracuse (\$12 million); Northeastern (\$13.5 million); Oregon Graduate Center (\$1 million); and Kansas (\$2 million).

Tim Beardsley

Falkland Islands

Opinions divided on penguin deaths

"NUCLEAR cargo" from the four British ships sunk during the Falklands conflict is polluting the South Atlantic, according to Soviet officials in Buenos Aires. At a press conference called last week to condemn President Reagan's unilateral repudiation of the SALT II agreement, the Soviet representatives claimed that the British military presence in the Falklands was a major threat to world peace, and that "the Argentines now have their own Chernobyl in the Atlantic".

This allegation was said to be based on research by the Buenos Aires correspondent of the Moscow weekly *Literaturnaya Gazeta*, Mr Vladimir Vesenskii, and on the Spanish newspaper *Cambio 16*. Vesenskii, the only civilian on the three-man Soviet panel, claimed that all four ships had had nuclear weapons on board and that *HMS Sheffield*, after being damaged by Argentine bombs, had been sunk by the British themselves "because of the danger of contamination". Vesenskii implied that contamination from these alleged bombs was responsible for the recent appearance of larger numbers of dead penguins "on the Malvinas beaches". Quoting an unnamed Argentine legislator, Vesenskii said that the "outer casings of the nuclear bombs were designed for dry places, not the ocean".

Asked if the Soviet Union had any information proving that there are now nuclear bombs on the Falklands, the Soviet charge d'affaires Viktor Tkachenko said simply that the Soviet Union and Argentina were "greatly concerned" about the building of a British military base on the islands, and called for Argentine–Soviet military relations to be strengthened at the "non-strategic level".

Vera Rich

● Results from UK laboratories seem to show that the penguin death-rate, in fact, is considerably worse than Vesenskii implied. He spoke of "tens" of dead penguins being washed up. In fact, from mid-February, elevated death rates of rockhopper penguins, and, to a lesser extent, gentoo penguins, began to be observed, and on 25 May, more than 3,000 dead rockhoppers were counted in a single rookery on New Island (off West Falkland).

Three batches of specimens have been sent to the United Kingdom, to the laboratories of the Ministry of Agriculture, Fisheries and Food (MAFF). Preliminary reports indicate that the birds died of starvation. Some appear to have suffered from puffinosis, a viral disease affecting sea-birds, and some show elevated lead levels in liver and kidneys. □

Strategic Defense Initiative

Britain wins first big contracts

BRITAIN has become the first foreign country to win a substantial contract for research in the US Strategic Defense Initiative (SDI). On Tuesday last week, George Younger, the Secretary of State for Defence, and Caspar Weinberger, his opposite number in the United States, signed a \$10 million letter of offer and acceptance in Washington. A further \$4.3 million contract goes to the United Kingdom Atomic Energy Authority's Culham Laboratory.

The contracts come at a time when there had been increasing criticism of British failure to obtain funds for research in the SDI programme. As Britain was the first country to back the US "star wars" project and to agree to cooperate on research and development, there were hopes that major contracts would be won — £1,000 million was announced as a possible figure last year by the then Secretary of State for Defence, Michael Heseltine. But there is now little hope that expenditure in Britain will come anywhere near this target.

The \$10 million contract is for a theoretical research study, expected to last 22 months, of the systems "necessary for the integrated defence of Europe". At present, that means the study will give a British view of the defence of Europe; at a later date, European allies may be added to the project. The contract has been put out to tender.

Culham Laboratory's contract is to cover research on neutral-beam particle systems. Neutral-beam weapons could form part of the first layer of an SDI system, shooting down ballistic missiles from space-based platforms. The Culham group, headed by Dr Tom Green, has already built up an international reputation for neutral-beam research. In collaboration with a French group, it designed the neutral-beam injectors for the Joint European Torus (JET) which sits alongside the Culham laboratories. The injectors fire intense beams of very energetic hydrogen and deuterium atoms into the tokamak to help heat up the plasma inside it. To generate the beam, positively-charged ions are accelerated electrostatically; they are subsequently neutralized in a gas cell, as charged particles would not be able to penetrate the intense magnetic field confining the plasma. Neutral beams are also essential for a space-based weapon as charged particles would be sent off course by the Earth's magnetic field. But for SDI application, where very high-energy beams are needed, negatively charged ions will be accelerated and an electron stripped off them before firing — it is here that the Culham group will have to develop new technology.

Any application for neutral beams as weapons is, of course, a long way off. Much basic research remains to be done. But if the present work, to extend over a three-year period, is successful, further SDI contracts may come along. The Culham group has already proved itself to be imaginative in finding research contracts; since the JET work drew to a close it has carried out research on semiconductor ion implantation, tokamak plasma beam diagnostics and ion thruster engines for keeping geostationary satellites on station.

Other SDI projects are rumoured to be in the pipeline for British research groups. The United States is interested in work on electromagnetic guns carried out at the

Royal Armament Research and Development Establishment and it is likely that a contract will be signed in the coming months.

Electromagnetic guns come in two varieties. In a rail gun two rails are bridged by an armature behind a metal projectile. When a current of several million amps is applied the expanding current loop fires the projectile.

Other electromagnetic guns are more similar to a linear induction motor; a series of switched coils induces a current in the projectile to fire it.

Either way velocities in the tens of kilometres per second can be achieved — a standard army rifle bullet goes at less than a kilometre a second. The aim is for velocities of 50–100 km per sec for electromagnetic guns which might be based in space or on land.

Alun Anderson

High-energy physics

Supercollider in the balance

Washington

ELEVEN US senators have signed a letter to President Reagan urging him to give his personal support to the Superconducting Supercollider (SSC), the 40 TeV proton-proton collider now at the design stage. Without strong White House support, the senators say the project "cannot go forward".

Already there are rumblings in Congress that the Challenger accident and the Gramm–Rudman deficit reduction act will make it hard to find the \$3,010 million needed to build SSC. The Department of Energy (DoE) steadfastly maintains that it has not decided SSC's future, but construction delays seem likely, and it is just possible that the project may be cancelled.

Senators Barry Goldwater and Dennis DeConcini, both Republicans from Arizona, organized the letter to President Reagan. While the letter describes SSC as "a very important national science project", the senators also find SSC attractive for the 7,000–8,000 jobs and millions of construction dollars it will bring to the state where it is built.

In testimony last week before the Senate Committee on Energy and Natural Resources, Alvin Trivelpiece, director of DoE's Office of Energy Research, called technical progress on SSC "excellent". But Trivelpiece said DoE is still weighing its options. A decision must be made soon so that the department's budget can reach the White House by September. Members of Congress in both the House and Senate feel a full-scale start for SSC in the 1988 budget is unlikely.

The House Appropriations Committee last month removed from the DoE budget request \$20 million that was to be used for continued development of SSC. While the committee's action must still be approved

by the full House, and later by the Senate, it was a signal that Congress is not interested in seeing more money put into SSC until its future is determined.

In May, DoE completed a favourable review of the conceptual design of SSC. The review described the \$3,010 million budget proposed for SSC as "credible and consistent with the scope of the project". The report also called the six-and-a-half-year construction schedule feasible, and said SSC will meet the requirements of high-energy physics "well into the next century".

But SSC's timing may be bad. William Graham still awaits confirmation by the Senate as director of the White House Office of Science and Technology Policy (OSTP). Until he comes on board as director, nobody at OSTP is anxious to carry the torch for such an ambitious project. The retirement of Hugh Loweth from the Office of Management and Budget, after 13 years as head of the Energy and Science Division, may also influence reception of the budget.

Senator Daniel Evans (Republican, Washington), a member of the Senate Energy and Natural Resources Committee, has declined to sign the Goldwater–DeConcini letter because he feels it would be irresponsible to commit funds to SSC when Congress is trying desperately to reduce the budget deficit. But he does have a modest proposal for DoE. Evans wants DoE to offer SSC as a "carrot" to the state it chooses for the site of the first deep geological repository for high-level nuclear waste. The three states chosen by DoE as final candidates for the waste storage site, Washington, Texas and Nevada, have so far been less than enthusiastic. All three have gone to court to try to block DoE's plans.

Joseph Palca

AIDS

Depressing news from Paris

Paris

THE razzmatazz of last week's international conference on acquired immune deficiency syndrome (AIDS) in Paris was tempered by gloomy facts and figures. Unless some effective form of therapy is found, 179,000 people in the United States will die of AIDS in the next five years; no doubts remain about the heterosexual transmission of the AIDS virus; at least half the members of several groups of intravenous drug users are now infected; the virus evolves as rapidly as the influenza virus; and it already has an alarming hold in some cities in central Africa.

There are still doubts about the validity of some of the African statistics, but those from Kinshasa, capital of Zaire, are copious and reliable. Of a group of prostitutes, 27 per cent have been infected by the virus as have 6.5 per cent of all the patients in hospital. In Rwanda, the equivalent fig-

ures are 88 per cent and 18 per cent, with 28 per cent recorded in pregnancy. With similar figures from other countries that border on Zaire, it is clear the virus is very well established and will continue to spread from mother to child and, most importantly, by heterosexual intercourse. That accounts for the approximately equal distribution of infection between the sexes in central Africa, whereas in developed countries male homosexuals still account for about 70 per cent of cases.

In western Africa the situation is less clear and, in many ways, more interesting. There is little, if any, evidence of the presence of HTLV-III/LAV, the standard AIDS virus, but increasing evidence of the distantly related HTLV-IV and LAV-2 viruses. It remains uncertain whether these are two names for the same virus. HTLV-IV was identified in Senegalese prostitutes whose health seems unaffected by the virus. As many as half the prostitutes in one unidentified (for political reasons) country in the region have been infected with HTLV-IV, with 38, 7 and 0 per cent in three nearby countries.

So far, not a single case of AIDS has been recorded in an HTLV-IV infected person, according to Max Essex of Harvard University. On the other hand, LAV-2 has now been isolated by the Pasteur Institute team led by Luc Montagnier from 8 AIDS or pre-AIDS patients from Guinea Bissau, Senegal and Cape Verde Islands, but its prevalence in prostitutes, or in general, is not yet known. There seems to be little prospect of an immediate exchange of material between Essex and Montagnier or, therefore, of establishing the precise relationship of the two viruses.

Hope remains that once the west African and monkey viruses have all been molecularly cloned and compared with HTLV-III/LAV, there will be a clear indication of which DNA sequences are responsible for pathogenicity and might form the best vaccines. Meanwhile, there is steady progress towards trials, perhaps later this year, of whether chimpanzees immunized with vaccinia virus engineered to contain the envelope gene of HTLV-III/LAV will be protected against infection with the virus.

There are, however, distinctly mixed views on the likely success of such a trial without (or even with) refinement of the envelope gene. Furthermore, there is emerging evidence of the rapid evolution of the virus within individuals. Thus Beatrice Hahn of the University of Alabama reported that consecutive isolates of the virus from individual patients had distinctive genetic maps, and comparable differences between the virus from infected blood donors and their recipients were

described by A. Srinivasan of the Centers for Disease Control in Atlanta. Probably both the direct evolution of the virus under pressure to escape the immune system and the emergence of different viruses from a mixed infecting population are involved. The chances of effective vaccination against such a moving target may be slim, as in the case of influenza virus.

For want of anything better as an advance in AIDS therapy, attention (certainly press attention) focused on the partial reconstitution of immune function in one AIDS patient given a bone marrow transplant from his healthy identical twin along with transfers of peripheral lymphocytes both before and after the transplant. But two other identical twins did not respond to the same form of treatment. Anthony Fauci, of the National Institutes of Health, intends to pursue such treatment, assessing the value of transplantation from matched siblings, the frequency of lymphocyte transfers and the value of simultaneous therapy with antivirals.

Little progress with antivirals was reported, although the first results of a double-blind placebo trial of azidothymidine are due this summer. Samuel Broder, of the National Cancer Institute in Maryland, who is persisting with the trial in the face of considerable pressure from some quarters simply to give the drug to all patients, is quietly optimistic about the outcome, particularly as the drug gains access at potentially therapeutic levels to the brain where it may reverse the loss of cognitive functions caused by the virus. Burroughs Wellcome, which makes azidothymidine, has exhausted the world supply of thymidine, from which it is made, but has plans to overcome this problem if the drug is shown to be valuable.

Without effective therapy, progression from infection to disease is not inevitable, but the cumulative incidence of AIDS in some well-studied cohorts of homosexuals is now around 30 per cent and showing no signs of decreasing with time. Moreover, Bob Redfield, of the Walter Reed Army Institute of Research in Washington, reported that 90 per cent of a series of patients who could be classified as belonging to one of 5 stages that precede full-blown AIDS progressed by at least one stage within 18–36 months.

Without an effective vaccine, the only means of prevention is to avoid intimate sexual contact with infected people, and to avoid their blood or blood products. There are clear signs that some homosexual communities have effectively taken this advice to heart. But it is less certain that drug users are doing so. In the worst reported case, 76 per cent of a group of Italian drug takers are now infected, compared with 6 per cent in 1980. In some groups fewer than 10 per cent are yet infected, but the reason for these differences is still unclear.

Peter Newmark

Japan's islands coming home

Tokyo

FOR years, the Japanese government has been doggedly trying to get the Soviet Union to return four small islands off the coast of Hokkaido seized in the closing days of the Second World War. Now a new survey shows that some of the islands will return quite soon — geologically speaking, that is.

The islands of Etorofu, Kunashiri, Shikotan and the Habomai group — the "northern territories" — were home to 16,000 Japanese before the Soviet invasion took their troops to an island within 5 km of the shores of Hokkaido in 1945. The continuing occupation is an endless source of friction between the two countries and has prevented Japan from concluding a post-war peace treaty with the Soviet Union, with all the trade benefits that might bring. But it seems the islands are creeping home at the rate of some 2 metres every 80 years.

The figures come from a resurvey of Hokkaido by the mapping division of the Geographical Survey Institute of the Ministry of Construction. A laser range-finder has revealed crustal movement that has occurred since an earlier triangulation survey carried out between 1908 and 1916. Throughout most of Hokkaido, net movement has been of the order of only a few centimetres or tens of centimetres. The two closest islands, the Habomais and Shikotan, are moving very much faster. If things continue as they are, they should collide with the mainland in about one million years.

David Swinbanks

European cooperation

Eureka gets down to business

EUREKA, the long-heralded but somewhat formless European programme of industrial cooperation on high-technology products, came a little more into shape on Monday, with the ministerial conference in London that marked the end of Britain's six-month secretaryship of the programme and the approval of over 60 new projects costing some 2,000 million ECU (£1,300 million).

Industry, it seems, is beginning to take Eureka seriously as a means of developing products — such as metal-working carbon dioxide lasers and advanced road vehicle guidance and communications systems — that can take a slice not only of Europe's market of 400 million consumers but also of the 120 million in Japan and the 240 million in the United States. Governments are involved because they can use their political weight to help remove trade barriers and develop common standards. Already five or six of the projects approved on Monday involve requests for additional political support of that kind, British officials say.

Monday's meeting, opened with evident enthusiasm by British Prime Minister Margaret Thatcher, approved a fivefold increase in the number of Eureka projects since the last ministerial conference in November, and established a significantly tiny secretariat of seven civil servants in — appropriately — Archimedes' Street in Brussels. The significance of the size of secretariat was that governments are determined not to create "another bureaucracy" but merely a clearing house for projects proposed by industry itself.

Research is not a problem. Europe's research laboratories "burst with talent and imagination", said Mrs Thatcher, but the old continent fell down in turning this into market success. To change that, Eureka is encouraging companies to share development. But Eureka "was not a source of funds" said Thatcher (this despite the efforts by some smaller countries and Italy to establish a central European funding system for the programme). Rather, Eureka seems set to develop with governments providing support only for their national companies within particular cooperations, a system which will clearly favour the larger countries.

The British government is setting aside no new money for Eureka, but, according to technology minister Geoffrey Pattie, it will find some £10 million a year from within the Department of Trade and Industry's "Support for Innovation" scheme to back the British companies in 28 Eureka projects. These projects, including company and other national support, will account for some £750 million of new technology development over the next

five years or so, Pattie indicated. But one-third of the Eureka projects involving Britain have no government support at all, said Pattie. West German government support is said to be on a similar scale. Italian research minister Luigi Granelli said on Monday that he "guaranteed" that every high quality proposal will get government support, but put no scale to funding which must finally pass a sceptical ministry of finance. Italian foreign minister Andreotti also hinted that Italian companies must beef up their proposals: there is no point in getting involved in Eureka "merely as a matter of prestige", he said. Meanwhile, the backing for Eureka in France, where the project originated 18 months ago, remains strong politically but uncertain financially, in a country which recently slashed its research and technology spending. Other countries, a total of 19 in all, including Iceland which joined on Monday, plan backing in proportion to the interest of their national companies.

Monday's meeting promised clarification

of one long-standing issue, the relationship of Eureka to the European Commission's own international research and development programme. According to Pattie, who for the next six months, which mark Britain's presidency of the European Council of Ministers, will chair the council of research ministers controlling the Commission's programme, "it is quite clear there will be a linkage between certain Eureka projects and the Commission's programme". According to Research Commissioner Karl-Heinz Narjes, the programmes can be complementary because the Commission's programmes are pre-competitive and Eureka's clearly competitive.

A senior commission official at the meeting said there could still be "considerable differences" in arranging cooperation between Eureka and Commission programmes, however. For example, the two large environmental projects (EUROMAR on sea pollution and EUROTRAC on tracing air pollution) included in Eureka on West Germany's insistence overlap strongly with Commission research and are not aimed at winning markets.

Robert Walgate

Laboratory safety

Just how hot is Harvard?

Boston, Massachusetts

HAVE the mistakes in the management of radioactive materials, for which Harvard University last month received a \$2,500 fine from the Nuclear Regulatory Commission (NRC) been corrected? Harvard vice-president Robert Scott says they have. But graduate students in the medical research area have a different tale to tell. Left unresolved, according to the students, all of whom were unwilling to be named, is the pervasive mishandling of radioactive materials stemming from short cuts taken because of research pressure and inadequate training.

"Safety regulations are often perceived as a waste of time, standing in the way of a Nobel prize", said one student. The risks of this approach are high — a laboratory stripped of its NRC licence loses its research edge as well as its personnel — yet a fear of a loss of prestige could cause management to look the other way. Dr Warren Wacker, director of Harvard's University Health Services, denies such blinking at Harvard. Using radioactivity is a privilege, he said, and he would brook "no nonsense and no pressure" in its management. As proof, he pointed to the three well-known laboratories in microbiology and AIDS (acquired immune deficiency syndrome) research that have just lost their licences for such infractions as inadequate cleaning of a spill.

NRC fined Harvard, according to an internal memorandum by its office of in-

spection and enforcement, less because of the twelve specific incidents cited than for "lack of effective management control and oversight" of radioactivity. As a result, the violations that occurred were of particular concern to NRC because "they could have resulted in unnecessary radiation exposure to individuals".

By law, NRC can fine nuclear power plants up to \$100,000 per infraction per day, but academic institutions only \$5,000. On a severity scale of 1–5, Harvard's infractions were rated as 3, for which the highest fine is \$2,500. Harvard has been warned about a number of other incidents, but they were not deemed serious enough to warrant a fine.

NRC, says Wacker, is cracking down on universities and Harvard, which is paying its fine without appeal, is not resisting. The salaries of two additional radiation health safety engineers will come from grants, not overhead. Other elements in Wacker's "model" programme will be continued — including surprise visits by Harvard's radiation service and immediate cancellations of licences for any serious infractions.

Radioactive materials are used on a daily basis in many medical research departments, amounting to "a couple of millicuries a week". Still, claims Wacker, "the highest level of radioactivity at Harvard is probably in the geology department in uranium ores, which are perfectly legal".

Elizabeth Collins

French science

Legal niceties end an era

FRENCH scientists were on the streets of Paris last week demonstrating against what placards described as the "strangling" of public research by a government that has cut ten per cent from its 1986 research budget, hitting research harder than any other field of government spending. After five heady years of strong support for science in France, laboratory staff and university lecturers are shocked, and were issuing posters last week claiming that French research was "in a coma".

One of the latest and most extraordinary blows, however, is not of the government's making. The Conseil d'État, effectively the French supreme court, has ruled that a body which for three years had made all the appointments and promotions in France's premier research council, the Centre National de la Recherche Scientifique (CNRS), was illegally elected.

This extraordinary decision, made on the basis of a legal technicality, to disestablish and annul the "Comité National" of CNRS, in theory invalidates all decisions taken by that body since 1983, thus affecting many of the academic scientists and technicians in France.

In practice, the decision may play into the government's hands, as its science advisers have long felt that the Comité National is too dominated by trades union interests. Paradoxically, it was a union that first made the legal complaint to the Conseil d'État back in 1983, over an election which had been designed to reduce union power.

Now, the ministry of research and the council of ministers will have the freedom to define a completely new election procedure at just the moment they would have desired — after a change of government from the political left to the right.

Appointments and promotions made by the Comité National since 1983 are thought likely to stand. But young researchers who recently attended interview boards arranged by the Comité and have been offered places in CNRS are less lucky. Some 375 of the 500 affected will be offered one-year temporary posts instead — to be reviewed next year when a new Comité is in place — and the others will fall by the wayside. This represents a 25 per cent reduction in the scientific recruitment rate projected by the previous government.

Robert Walgate

Italian biotechnology

New initiative is called for

Rome

ITALY needs to spend a million million lire (£420 million) on a massive biotechnology programme over the next five years, according to a new report from the Italian National Committee for Biotechnologies. That, says Minister of Research Luigi Granelli, will give the government a tool to make Italy competitive in previously neglected sectors such as chemicals and agriculture.

Four areas are earmarked for funds: fundamental research in public centres and universities including three special projects from the Consiglio Nazionale della Ricerca (CNR), the national research council (25 per cent of funds), national research involving both scientific bodies and industry (40 per cent), the finance of research societies and stimulation of industrial activity (25 per cent) and training of researchers (10 per cent).

Established in July 1985 by Granelli, the committee includes 14 representatives from universities and public research centres and eight from industry. Coordinated by Professor Arturo Falaschi, director of the CNR Institute of Genetic Engineering in Pavia, the report looks largely at the short term. In countries such as the United States and Japan, says the report, basic research in biotechnology is

strongly supported by government, with private investment strong in biomedicine in the United States and in agriculture in Japan. In Europe, adds the report, the United Kingdom leads in biotechnology, but good state-financed programmes have been launched elsewhere.

Italy has good scientists, says Professor Luigi Rossi Bernardi, CNR president, as well as a good research tradition at the Institute of Molecular Biology. And there is the new International Centre for Genetic Engineering and Biotechnology in Trieste (see p.3). But this is not enough for Italy to catch up with other nations.

The biomedical field, says the report, is likely to yield returns in the short term, while agriculture, the food industry and fine chemicals have good medium-term prospects. So the report suggests doubling in five years the number of researchers in biotechnology in public centres such as CNR and ENEA (the national centre for alternative energy), as well as in the universities and health ministry, rather than in industry.

To coordinate the programme, it is suggested that either the present committee should be institutionalized or a National Institute of Biotechnology should be set up like that for nuclear physics.

Paola di Paoli

US space

Air Force moves on launchers

Washington

THE US Air Force has announced plans to build at least twelve medium launch vehicles (MLVs) as an alternative to the space shuttle for launching military payloads. In seeking the new rocket, the Air Force has indicated it wants potential suppliers to develop a commercial version of the rocket to compete with Europe's Ariane for commercial payloads.

The new rocket will be able to launch 10,000-lb payloads into low Earth orbit, and 2,200-lb payloads into a 10,000-mile high circular orbit. The Air Force intends to use the new rocket primarily to launch its Navstar satellites, part of a new Global Positioning System under development.

Air Force interest in a complementary commercial vehicle may seem odd, but building a launch vehicle with commercial applications would have the obvious benefit of spreading development costs. Air Force Secretary Edward C. Aldridge has consistently shown concern for the viability of the US space programme, according to space policy consultant Christopher Roberts, and by providing a steady customer, the Air Force can go a long way towards assuring the success of a commercial launcher. John Pike of the Federation of American Scientists takes a more cynical view, believing that the new MLV results from an "unholy alliance" between the Air Force and the rocket manufacturers. As Pike sees it, the Air Force gains more positions for uniformed officers in its space command, and the manufacturers get to crank up their production lines.

Money for starting the procurement process of the new MLVs was in the urgent supplement appropriations bill passed by Congress last week. Also in the appropriations bill were funds for an additional 13 Titan 34D7 rockets, the workhorse of the Titan series. With the new orders, Martin Marietta, the manufacturer of the Titan 34D7, will be in a position to compete with Ariane-4, due to launch next year, should the company decide to enter the commercial launch arena.

In a related development, the National Aeronautics and Space Administration (NASA) announced last week that it was offering Indonesia the opportunity to launch its Palapa communications satellite aboard a Delta rocket. The Palapa satellite had been scheduled to fly on the shuttle last month. If Indonesia accepts, NASA will have to build one more Delta to meet its launch commitments. But NASA denies that the new Delta would mark NASA return to expendable launch vehicles.

Joseph Palca

UK science

Save British Science takes off

AN extra £100 million pounds a year is the minimum needed to "save British science" according to the new pressure group of the same name. The figure comes in the first major effort by Save British Science (SBS) to win influence in Parliament: the presentation of evidence to the House of Lords Select Committee on Science and Technology.

The SBS campaign was born last January with a half-page advertisement in *The Times* newspaper protesting against the underfunding of UK scientific research. At that time, it was not a formal organization, but the response the advertisement provoked made it seem worthwhile to become so. More than 2,500 scientists have already contributed financially. From last week, SBS is a formally constituted body and will start to collect membership fees from individual scientists and students and from societies and companies. Its aims are to "communicate to the public, parliament and the government a proper appreciation of the economic and cultural benefits of scientists' research . . .". A London office is being opened.

Individual members of SBS have argued their case with the recently departed Secretary of State for Education and Science, Sir Keith Joseph, and have urged scientists to write to their Members of Parliament, but the evidence to the House of Lords is their first major move.

The House of Lords select committee has just finished taking evidence on the state of British science — SBS actually squeezed in just after the deadline. The committee is expected to send a report to the Secretary of State or to call for a debate in the House of Lords on the state of science, perhaps in the next session.

SBS's evidence is not an overall analysis of British science but a collection of real-life stories from scientists depicting the difficulties they face. Taken as a whole, it should depress their lordships: libraries where core international journals can no longer be afforded, electron microscopes and other vital equipment without funds to keep them in use, talented post-doctoral students who vanish abroad, equipment that ceased to be up to date a decade ago, time wasted on an endless struggle to obtain funds, and out of work young researchers who carry on as "DHSS fellows" (living on social security payments).

A case is made that the damage extends to research with clear economic potential. Experts in high growth areas such as electronics are packing up and going abroad where they will help to train the next generation of foreign competitors. The fear is expressed that once Britain falls too far

behind, it will become impossible to catch up again and there will be nothing to offer in international exchange and collaboration. Evidence is given that Britain is already regarded as not being a serious partner in European ventures.

For Save British Science, the remedy lies in tackling the "alphas", grant proposals given the highest priority rating. In the past, these would have received automatic support. Now, around three-quarters are funded to about two-thirds of the sum requested, but full amounts are rarely asked for as there is so little chance of success. Adequate support for all alpha-graded proposals plus the basic infrastructure needed for them would cost an extra £100 million a year.

That money would still, however, be

little more than a stopgap. What SBS wants is a coherent scientific policy that would raise the low level of investment in research by UK industry in comparison with other nations and establish a "new sharing of responsibilities between government, higher education and industry".

The future success of SBS's own campaign must depend on its forming close links with industry. The only argument that is likely to carry real weight with the government is that basic research is essential if there is to be applied research that benefits British industry.

So far, SBS's support has come almost entirely from academics, implying that the two constituencies of industrial and university researchers are far apart. In the future there will be efforts to recruit industrial support, beginning with companies that have links to universities and those dependent on high technology.

Alun Anderson

Ocean satellites

Topex launch comes closer

Washington

CONGRESS has given the National Aeronautics and Space Administration (NASA) a green light to seek final design proposals for the Topex/Poseidon satellite. A joint project with the French space agency (CNES), Topex/Poseidon will provide the most accurate data on ocean surface topography ever recorded.

Topex/Poseidon is scheduled for launch in 1991 aboard the French rocket launcher Ariane. Three companies are in the running for the final contract to build the satellite: Fairchild Industries, RCA and Rockwell International have all completed design proposals, and will now provide a final design to NASA by the end of September. In addition to approving selection of the hardware, Congress is allowing NASA to solicit proposals from potential users of the satellite. According to the agreement with CNES, the French will review scientific proposals from most of Europe and Africa, while NASA will judge proposals from the rest of the world.

One of the primary beneficiaries of the ocean topography data will be the World Ocean Circulation Experiment (WOCE). Carefully calibrated altimetry data from Topex/Poseidon will provide an accurate picture of ocean currents. Because Topex/Poseidon has a design life of three years — and will carry enough fuel to last two years longer — oceanographers expect to build up a picture of variations in ocean currents as well as their steady-state behaviour.

Both Japan and the European Space Agency (ESA) also have plans for complementary satellites. Japan's Marine Observations Satellite (MOS-1) and ESA's Earth Remote Sensing Satellite (ERS-1) will carry altimetry instruments.

Having several satellites in orbit concurrently will provide useful cross checks on their data, says Worth Nowlin of Texas A&M University, co-chairman of the US WOCE scientific steering committee. Nowlin also says the accuracy of Topex/Poseidon data will help to validate less precise measures from the other satellites. Another advantage of Topex/Poseidon will be its orbit. At 63.1 degrees-inclination and 1,334-km altitude, Topex/Poseidon will pass over the same point on Earth once every ten days. Using a laser tracking system, NASA and CNES will be able to calibrate the satellite's altimeters regularly. In addition, because it will not be launched into a Sun-synchronous orbit, Topex/Poseidon will provide data on solar tides that typically confound ocean current measurements.

Topex/Poseidon fits nicely into the scheme of coordinated Earth observations foreseen by the Earth System Sciences Committee in its report to NASA released last week (see *Nature* 321, 801; 1986). The committee has endorsed the Geopotential Research Mission (GRM) as a sequel to Topex/Poseidon. GRM will provide detailed measurements of the Earth's gravitational field, giving an absolute reference point for sea-surface heights measured by Topex/Poseidon.

Money for the final design stage of Topex/Poseidon has already been allocated to NASA, but Congress must still approve funds for full-scale development. Until the Reagan administration says how it plans to balance manned and unmanned missions in the wake of the Challenger accident, Congress is likely to postpone definite decisions about support for NASA projects.

Joseph Palca

Darwin, Lyell and gradualism

SIR—In an elegant essay, Rhodes¹ convincingly demonstrated that Darwin was not a strict gradualist — that he was, in fact, a pluralist and his views overlapped what over the past decade or so has been hailed as a new theory, “punctuated equilibria”. According to Rhodes, Darwin recognized a component of stasis, a tenet of punctuated equilibria, and repudiated the gradualistic aspect of evenness in transformation but “...whether it was ‘slow’ depends on our definition of speed”; slowness is not synonymous with gradualism. Darwin, Rhodes correctly points out, clearly appreciated the significance of local, periodic speciation and its representation in the fossil record. Although rapidity of speciation is associated with punctuation, Rhodes argues that this rapidity might well be encompassed within the “gradualism” of Darwin; whether “gradual” or “rapid” becomes a matter of definition.

Recently, Dawkins², in his review of a book *Time Frames: The Rethinking of Darwinian Evolution and the Theory of Punctuated Equilibria*³ by one of the originators of the theory, Niles Eldredge, substantiates many of the claims of Rhodes. Although the fervent punctuationists have stressed the revolutionary and anti-darwinian aspects of the concept, Dawkins states that “it lies firmly within the neo-darwinian synthesis”. Darwin emphasized the term “gradual” in his campaign against creationism; in such a context, explains Dawkins, gradualism is practically synonymous with evolution itself. “The punctuationists deluded themselves that they were saying something revolutionary. And it was all due to a simple verbal misunderstanding: a confusion of two senses of the word ‘gradual’.”

In a very recent defence of “lyellian uniformitarianism”, I⁴ have, by referring to Lyell's original works, demonstrated that he certainly appreciated episodicity, both past and present, and clearly incorporated that concept in his use of “gradual”. In a parallel situation (and with some overlap) to Darwin's combat with creationists in the palaeontological and biological realm, Lyell was waging an intense philosophical battle against the catastrophists, a still very vociferous and influential group in the 1830s when his early editions were written. In comparison with this antagonistic philosophy, which emphasized truly cataclysmic events beyond our present knowledge, Lyell's uniformitarianism seemed gradualistic, and was — relatively speaking; he used “gradual” in a broader sense than most have realized. One quotation from the first edition of his *Principles of Geology* describing the episodic uplift of a fault-bounded mountain range serves as a rep-

resentative statement of his concept of “gradual”: “We know that one earthquake may raise the coast of Chili for a hundred miles to the average height of about five feet. A repetition of two thousand shocks of equal violence might produce a mountain chain one hundred miles long and ten thousand feet high. Now, should only one of these convulsions happen in a century, it would be consistent with the order of events experienced by the Chilians from the earliest times⁵.”

In another recent review, Cloud⁶ criticized authors who labelled Lyell a gradualist, advising them to review original sources. Cloud has suggested (personal communication), correctly I feel, that Lyell's sense of gradual and uniform was with regard to the “long average”. Although he frequently used the word uniformity, it seems clear from content that what he had in mind was far from simple gradualism.

I think that both Darwin and Lyell have been rather seriously misjudged with regard to their use of “gradual”. For more than a century, it seems we have cast aspersions on their gradualism without appreciating the context of their times and, apparently, frequently without carefully consulting the original sources.

DONALD H. ZENGER

Department of Geology,
Pomona College,
Claremont, California 91711, USA

1. Rhodes, F.H.T. *Nature* **305**, 269–272 (1983).
2. Dawkins, R. *Nature* **316**, 683–684 (1985).
3. Eldredge, N. *Time Frames: The Rethinking of Darwinian Evolution and the Theory of Punctuated Equilibria* (Simon and Schuster, New York, 1985).
4. Zenger, D.H.J. *J. Geol. Ed.* **34**, 10–13 (1986).
5. Lyell, C. *Principles of Geology* Vol. 1 (John Murray, London, 1830).
6. Cloud, P. *Bull. Am. Ass. Petr. Geol.* **69**, 138–139 (1985).

Problems of Africa

SIR—I have often wished that your Opinion and News columns showed more awareness of the Third World's worsening problems; but if your editorial “Who will pity Africa?” (*Nature* **321**, 548; 1986) is the best you can do, it might be better to remain silent. In one column of pontification you reveal a remarkable ignorance of Africa's history and economy, and even of recent events in the continent.

You say that Uganda, once a kind of paradise, “has now settled for inter-tribal violence and the chaos that follows”. Apart from the fact that Britain fostered tribal rivalries in colonial days by recognizing the Kabaka of Buganda as ruler, and then backed two disastrous leaders (Obote, Amin and Obote again) who deepened these divisions, you seem to have missed the news earlier this year that Uganda now has a widely popular president (Museveni) who offers real hope of

reconciliation.

Your reference to Nigeria having “squandered” its oil wealth is also somewhat misplaced, in view of the fact that Britain's oil boom appears to have financed nothing more tangible than tax cuts for the better-off. Would you call this a good example?

Worse than all of this, however, is your failure to comment on the external factors, particularly falling commodity prices, that lie at the heart of Africa's crisis. For example, Tanzania spends 60 per cent of its external earnings on oil, and while 1 tonne of tea paid for 60 barrels of oil in 1972, by 1980 it bought only 4½ barrels. In the same period, the amount of coffee needed to pay for a lorry rose from 5 to 28 tonnes. This is the “free market” over which Africa has no control.

There has been no shortage of “good ideas” on what to do next, but Western governments have found most of them deeply unpalatable. Instead of joining Mr Schulz in lecturing Africans on their improvidence, you could start by asking why Britain and the United States have so far failed to contribute to the special programme for sub-Saharan Africa set up by the International Fund for Agricultural Development.

MICHAEL SPENCER

34 Bayham Road,
Sevenoaks,
Kent TN13 3XE, UK

Condoms in Japan

SIR—Legalization of the contraceptive pill seems to be regarded as an entirely positive development. The cautiousness of the Japanese authorities has, however, contributed greatly to the wide popularity of the condom, which is more frequently used than any other form of contraception.

The low incidence of acquired immune deficiency syndrome (AIDS) in Japan and the relaxed attitude towards the possible spread of the disease among the general population, as well as the low incidence of infertility caused by salpingitis, may be explained in part by the widespread use of the condom.

Rather than Japan copying Western habits of contraception, it would be better for those in the West to attempt to understand (and possibly copy) the cultural patterns that make this simple contraceptive acceptable to Japanese men and women rather than being considered unnatural and unwelcome.

BO WARMING

Sperm-Help Foundation,
Bjelkes Alle 46,
DK-2200 Copenhagen N,
Denmark

1. *Nature* **317**, 760 (1985).
2. Coleman, S. *Family Planning in Japanese Society* (Princeton University Press, 1983).
3. *Nature* **318**, 306 (1985).

The proper study of mankind

Molecular biology has made the human genome accessible to laboratory investigation. But does that mean that the sequence of the human genome would be worth the effort?

THERE is no scientific reason for studying man. Thus one of the 120 speakers at this year's Cold Spring Harbor Symposium on the molecular biology of *Homo sapiens* from 28 May to 4 June. The issue arises because, as the programme was designed to demonstrate, there are now few fundamental questions in biology that cannot be explored using human genes and human cells.

This in itself is scarcely sufficient reason for overlooking the rather obvious shortcomings of man as an object of investigation. But there is one reason, just the same: the human species alone preserves its rare defective variants. Mouse mothers, whose matings are manipulable and whose generation times are manageable, tend to eat offspring that seem not quite right before they have grown large enough to be investigated, and thus destroy the rich mine of information that medical genetics has made available in man.

If we accept that man's sole unique contribution is his genetic diseases, the content of the Cold Spring Harbor Symposium has some interesting reflections to offer in the light of the debate, discussed in *Nature* (321, 371; 1986) and formally tabled at the symposium, on whether to launch into the gargantuan exercise of sequencing the entire human genome.

There have been two outstanding contributions of medical genetics to science. One is the analysis of mutant haemoglobins; the other is the remarkable achievement of Michael Brown and Joseph Goldstein (University of Texas, Dallas) who uncovered the cellular and molecular basis of the hypercholesterolaemia that predisposes to heart disease (see *Nature* 317, 569; 1985), the essence of which was reviewed by Goldstein at Cold Spring Harbor.

Cholesterol is cleared from the bloodstream in the form of a lipid-protein complex (LDL) that binds to a specific receptor (the LDL receptor) on the surface of cells, whence it is internalized by endocytosis and released to provide the substrate for the synthesis of cell membranes and steroid hormones, as appropriate. If the receptor-mediated internalization fails for any reason, the cholesterol builds up in the bloodstream, atherosclerotic plaques develop and heart disease ensues. By investigation of precisely why internalization does fail in individuals with familial hypercholesterolaemia, Brown

and Goldstein have built up a picture of the cellular and, more recently, the molecular biology of receptor-mediated endocytosis which has provided fundamental insights into these processes as well as clarifying the mechanism of atherosclerosis.

This has been possible only because Brown and Goldstein worked up to the molecular biology from classical genetics via biochemistry. More characteristic of latter-day triumphs were the papers presented by Louis Kunkel (Boston Children's Hospital) and Stuart Orkin (Harvard Medical School), both published in this issue of *Nature* (pp 73 and 32). They have arrived directly at the genes responsible for two human genetic diseases without engaging with the intervening biochemistry.

Kunkel and his collaborators have succeeded in forging a path to the gene for Duchenne muscular dystrophy (DMD); Orkin and his colleagues have identified, cloned and sequenced the gene for X-linked chronic granulomatous disease (CGD). Neither group is any the wiser about the molecular or cellular mechanisms of the disease.

From this point of view, DMD and CGD present problems of a different order. DMD is a programmed wasting disease of muscle whose mechanism and site of action are wholly unknown, whereas CGD is an immune deficiency known to be consequent upon a defect in the oxidase system in phagocytic cells. Peter Goodfellow explains in an article on p. 12 the route by which Kunkel and his collaborators have arrived at a region of 20,000 base pairs comprising possibly no more than a part of the DMD gene, and what this implies for the daunting prospect of identifying the gene product or products. The problem here, and much of the excitement, is that there is so little idea what to look for.

The CGD gene is much more tractable. Starting with genetic linkage analysis using polymorphic DNA probes, Orkin and co-workers identified a sequence that is transcribed specifically in normal phagocyte but not in the phagocytic cells of three of their four CGD patients. The fourth patient produced a transcript containing a small deletion that presumably abrogates the production of a functional protein. Whether the transcript encodes one of a number of known but poorly characterized components of the oxidase system, or some so far unknown enzyme,

may require some ingenuity, as well as some biochemistry, to discover.

The substantial challenge now facing Orkin and Kunkel and their collaborators is in no sense a reflection on the quality of the work or the validity of their approach. But it is an inevitable consequence of approaching the DNA directly, and thus has a direct bearing on the desirability of sequencing the rest of the human genome.

In the discussion on that topic at the symposium, Paul Berg (Stanford) set out the issues — is it feasible, who will pay and is it worth it? Walter Gilbert (Harvard) seraphically chalked up the tally — three thousand million bases at 10^5 bases per year equals 30,000 person-years or, at the current rate of 2×10^6 bases per year, but allowing for the 6 million already sequenced, 1,000–1,500 years to finish all of it; at a going rate of \$1.00 a base.

Nobody seems to doubt that it is feasible; nor that with improvements in technology it can be done in perhaps a century or less. The question is whether it is worthwhile. There is a general conviction that, even with funding from the US Department of Energy (see *Nature* 321, 371; 1986), the project would be bound to divert resources from other areas of research. In return, it is argued, it will encourage innovative technology and provide an information resource that would nourish other areas.

But technical innovations are already occurring with impressive momentum; and as an information resource, the sequence of the human genome is an extremely doubtful asset. If the skill and ingenuity of modern biology are already stretched to interpret sequences of known importance, such as those of the DMD and CGD genes, what possible use could be made of more sequences? Moreover, it is believed that roughly half the human genome is nonsense and repetition; and those who argue that there may be meaning in the nonsense and function in the repeats are not likely to prove their point by sequencing it all.

The difficulty is the same for all attempts to answer biological problems by reading DNA: unless you have a very good idea how to phrase the question, the sequence is not going to give you the answer. Blind sequencing at the expense of good ideas would, in the words of David Botstein, be for biologists to "indenture [themselves] to mindlessness".

Miranda Robertson

Duchenne muscular dystrophy

Collaboration and progress

from Peter N. Goodfellow

PROGRESS in science often reflects the Hegelian dialectic between collaboration and competition. Collaboration allows the most efficient use of resources and competition can generate original approaches to solving problems. Geneticists, molecular biologists and clinicians studying Duchenne muscular dystrophy (DMD) have demonstrated both originality and a willingness to collaborate. The scale of that collaboration can be estimated from the 75 co-signatories to the paper¹ on page 73 of this issue.

DMD, a devastating X-linked genetic disease afflicting one in every 4,000 newborn boys, is characterized by progressive muscle wasting that is fatal during the second or third decade of life. The apparently simple X-linked mode of inheritance makes DMD a model system for the application of the techniques of molecular genetics. For any genetic disease of unknown aetiology the strategy is the same (see ref. 2): pinpoint the location of the defective gene; starting from neighbouring reference points 'walk' to the gene by isolating overlapping DNA sequences; and deduce the nature of the disease from the sequence of the gene and the gene alterations found in disease sufferers. For DMD the gene has been localized and chromosomal walks have begun. Unfortunately, nature has a habit of putting new twists into the theoretical constructs of scientists. New results reviewed at the recent Cold Spring Harbor Symposium (see page 11) and described by L. Kunkel and collaborators³ suggest that DMD gene identification and description will prove difficult: either the gene is very large or the region is very complex.

Three different paths have led to the chromosomal position of the DMD gene. The first followed from the discovery of rare female DMD patients resulting from chromosomal translocations between one X chromosome and an autosome. In each case the breakpoint on the autosome varies but the breakpoint on the X chromosome is similar^{3,4}. In normal females one of the two X chromosomes is inactivated to avoid gene-dosage problems. What happens in the DMD females is that the normal X chromosome is inactivated, and DMD results because the breakpoint on the translocated X chromosome disrupts the DMD gene. The second path follows from classical mendelian genetics and involves the identification of sequences genetically linked to the DMD gene. The third path followed from the discovery of a boy with DMD who had a deletion of DNA from within the short

arm of the X chromosomes⁵. None of the paths was well signposted and there was potential for straying. But the fact that all three paths led to the same point within Xp21 on the short arm of the X chromosome provided overwhelming evidence that this was the chromosomal position of the DMD gene⁷.

The next step was to isolate the DNA sequences within the region Xp21 that correspond to the disease gene. Unfortunately, this region contains about 10⁷ base pairs (bp) of DNA and the closest DNA sequences used as genetic markers were at least 10 per cent recombination away from the DMD gene. If it is assumed that 1 per cent recombination corresponds to about 10⁶ bp, walking the 10⁷ bp to the disease locus would be a daunting task. In the near future, the new technologies of chromosome jumping and hopping may ameliorate the worst aspects of chromosome walking (see the recent *News and Views* article⁸ by Peter Little). But the groups of L. Kunkel (Boston) and R. Worton (Toronto) decided not to wait for the future.

The short-cut taken by Kunkel and colleagues was to develop and exploit an accelerated hybridization technique which enabled them to clone, by subtraction, sequences defined by chromosomal deletions^{8,9}. That taken by Worton and colleagues was to clone the sequences at the breakpoint present on the X chromosome of a female DMD patient¹⁰. One of

the sequences described by Kunkel and collaborators is deleted from between 5 and 10 per cent of all DMD patients. In several other human genes about 10 per cent of mutations causing disease are caused by small deletions within the gene (for example, *HPRT*; ref. 11). By analogy with these genes it was thought that the sequence *DXS164* was derived from within the DMD gene. It was therefore very disappointing to learn that 5 per cent recombination occurs between *DXS164* and the DMD gene; that is, if *DXS164* is part of the disease gene it would not be expected to recombine with the disease phenotype at a detectable frequency. Another disappointment was the discovery that the sequence XJ-1.1 isolated by Worton *et al.* from the X-chromosome translocation breakpoint also recombines with the DMD gene at about 5 per cent despite the fact that some patients had deletions which included both XJ-1.1 and *DXS164*. Perhaps these observations were not totally unexpected — careful cytogenetic observation had suggested that not all the X-chromosome breakpoints associated with DMD were precisely in the same position⁴ and chromosomal translocations may cause deleterious effects over large areas.

In an attempt to regain the path leading directly to the DMD defect, Kunkel *et al.* describe in this issue the cloning of 140,000 bp of DNA around the *DXS164* locus and the use of the isolated sequences to analyse DNA from 57 DMD patients with deletions. As the authors point out, ascertainment of the deletions may be biased because more patients have been screened with the original *DXS164* probes than have been screened with probes from

100 years ago

The recent discoveries at Tiryns

THE excavations, made during the last two years at Tiryns, by Dr. Schliemann and Dr. Dörpfeld, have thrown new light on what has been hitherto an almost unknown period of Greek history when Hellenic civilisation had not yet emerged from its Oriental cradle, nor developed its highly cultured systems of social and political government.

The literature of Greece has made us familiar with the later times, when the individual was for the most part merged in the State, and when the wealth and artistic skill of each city was devoted to public uses, such as the Council-chamber, the Agora, or the stately temples of the gods, rather than to the luxury of any one person.

But at Tiryns a very different picture is presented to us: we see a single autocratic chieftain, ruling in a sort of feudal state, and occupying a magnificent palace, surrounded by the humbler dwellings of his circle of retainers; while, instead of the utmost resources of the architect, the sculptor, and the painter being lavished on the shrine of the presiding deity, a mere open-air altar is dedicated to the



PLAN OF THE PALACE MEASURED BY DR. DÖRPFELD.

god, and it is the chieftain's house which is decked out with the splendours of gilt bronze, marble sculpture, and painted walls.

From *Nature* 34, 218; 8 July, 1886.

other parts of Xp21. Most of the deletions have removed the whole 140,000 bp of the expanded *DXS164* region. One deletion of 40,000 bp is completely contained within the *DXS164* region and most of the other deletions overlap each other in the *DXS164* region. However, five deletions within *DXS164* show no region of overlap. The breakpoints in the latter deletions may be separated by up to 80,000 bp of DNA. The only obvious conclusion is that the deletions are large, a conclusion supported by the deletion of the flanking markers and, occasionally, adjacent disease loci. These deletions are quite unlike the small deletions associated with mutations at other disease genes, although a similar spectrum of large deletions on the Y chromosome has been associated with the generation of XY females¹².

Conclusions about the nature of the *DMD* gene must remain speculative and are based on two plausible, but unproven, assumptions. The first is that the deletions are simple and not part of complex rearrangements, which can be tested. The second assumption is that the deletions, when present, are responsible for the disease and are not merely markers for the disease. Accepting these limitations, the simplest interpretation is that either the *DMD* gene is large or mutation in several different genes in the same region can cause the disease. The deletion data are consistent with a single large gene of 100,000–200,000 bp — no larger than the *factor 8C* gene¹³. Less consistent with this view is the apparent variable position of the *DMD* locus as defined by the recombination and deletion studies with respect to the *DXS164* locus. To accommodate these results either one or several genes need to be spread over a distance as large as several million base pairs. Alternatively, recombination rates in this region may be greatly elevated. Neither possibility is without precedent: the bithorax gene of *Drosophila* is greater than 100,000 bp, the same fraction of the human gene would represent 2,000,000 bp; and in human male meiosis the pseudoautosomal region shows a 10-fold elevation in recombination.

A more radical explanation of the results acknowledges the variable order of *DMD* with respect to *DXS164*. If a polymorphism exists in the human population for inversion of a few million base pairs in the Xp21 region, several of the observed phenomena could be explained. Regions of inversion would lead to different gene orders in different families. Recombination outside the inversion followed by a second abnormal and rare (10^{-3}) recombination event within the inversion (perhaps between ubiquitous and inverted repeat sequences) would result in the generation of large deletions. This model predicts the concomitant generation of duplications and the observed high frequency of mu-

tation at the *DMD* locus. A disadvantage of the model is that it requires the new assumption of an inversion present in the population at high frequency. Molecular analysis of deletions of the gene encoding the LDL receptor demonstrates repeat sequences flanking the breakpoints within the gene (D. Russel, Dallas). However, Kunkel *et al.* have not seen a similar preponderance of repeats associated with breakpoints in *DMD*.

Several groups are attempting to resolve these questions by constructing restriction maps of the Xp21 region of the X chromosome using pulse-field gradient gel electrophoresis (K. Davies, Oxford; P. Pearson, Leiden; see ref. 2). But even more pressing is the need to find transcribed sequences within the *DXS164* region. Several strategies are possible, but the most straightforward is to screen messenger RNA isolated from adult muscle with probes derived from the expanded

DXS164 region. Another approach is to screen for conserved sequences between man and mouse. Using this method Kunkel *et al.* find potential exons within the *DXS164* locus which are also X-linked in the mouse and Worton and collaborators find conserved exons near XJ-1.1. □

1. Kunkel, L.M. *et al.* *Nature* **321**, 73 (1986).
2. Little, P. *Nature News and Views* **321**, 558 (1986).
3. Lindenbaum, R.H. *et al.* *J. med. Genet.* **16**, 384 (1979).
4. Boyd, Y. & Bucke, V.J. *Clin. Genet.* **29**, 108 (1986).
5. Davies, K. *et al.* *Nucleic Acids Res.* **11**, 2303 (1983).
6. Francke, U. *et al.* *Am. J. hum. Genet.* **37**, 250 (1985).
7. Craig, I.W. & Goodfellow, P.N. *Lab. Invest.* **54**, 241 (1986).
8. Kunkel, L.M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 4778 (1985).
9. Monaco, A.P. *et al.* *Nature* **316**, 842 (1985).
10. Ray, P.N. *et al.* *Nature* **318**, 672 (1985).
11. Yang, T.P. *et al.* *Nature* **310**, 412 (1984).
12. Vergnaud, G. *et al.* *Am. J. hum. Genet.* **38**, 109 (1986).
13. Gitschier *et al.* *Nature* **312**, 326 (1984).

Peter N. Goodfellow is in the Laboratory of Human Molecular Genetics at the Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK.

Geomagnetic reversals

Evidence for asymmetry and fluctuation

from Jeremy Bloxham

FIELD reversals, occurring roughly every million years, are the most dramatic of the wide range of phenomena exhibited by Earth's magnetic field. And the next reversal on Earth may not be so far away: if the current rate of decay of Earth's dipole component is maintained it will vanish in less than 2,000 years' time. Reversals are studied using palaeomagnetic techniques to recover the remanent magnetization of lavas and sediments; in recent years, attention has focused on obtaining reliable records spanning actual reversals, in order to build up a description of the field in transition. Two papers in this issue^{1,2} present such records obtained from sediments.

Most previous work has concentrated on using lavas to study the field in transition. From these studies a picture, albeit rather blurred, has begun to emerge as to how the field reverses, and it has been possible to draw some tentative conclusions about the working of the dynamo during a reversal.

First came the recognition that the reversing field is non-dipolar³. The more interesting issue of whether the field is axisymmetric is still undecided: the evidence tends to suggest that reversals are initially axisymmetric, but may become non-axisymmetric as the reversal progresses. Hide suggested⁴ that reversals are initiated if the field becomes axisymmetric, since axisymmetric fields cannot be maintained by dynamo action and must decay

diffusively. However, it appears that reversals are not entirely diffusive; instead, fluid motions in the core actively contribute to the reversal process. This viewpoint is supported by the observation of Prévot *et al.*⁵ of very rapid variations in the field during reversals, variations with characteristic times very much shorter than appropriate diffusive timescales. Other work has indicated that reversals take place in two distinct phases^{6,7}; in the first phase an intermediate state is reached from which either the reversal is completed, or the field returns to its original polarity.

On page 27 of this issue Valet *et al.*¹ report results obtained from marine clays from western Crete. They find evidence of short-term variations in the field, the amplitude of which seem to increase during the reversal. They also find evidence of an intermediate state. Now that studies based on both lavas and sediments have indicated some of the same characteristics of reversals, the evidence that such characteristics are real is greatly strengthened. In a forthcoming issue, Herrero-Bervera and Thayer² will report on four transition records obtained from deep-sea sediments from the north central Pacific Ocean. They find evidence that the transition field is non-axisymmetric and that it is characterized by quite rapid fluctuations. They also consider which of the two common phenomenological models of the reversal process is favoured by their data. The

flooding model, whereby the reversal is believed to be initiated in some region of the core and subsequently spreads, or floods, throughout the core, seems to be favoured over the standing model, whereby the dipole reverses while most of the non-dipole field is unaffected.

Valet *et al.*'s results also indicate the presence of longer-term variations, of between 2,000 and 4,500 years periodicity. Sediments are a very good way of detecting such long-term variations, although it is hard to pinpoint their precise period because of uncertainties in the sedimentation rate. Interestingly, these variations appear to be unaffected by the reversal. They claim that this observation has some relevance for the validity of the frozen-flux approximation.

The frozen-flux approximation — which has recently become the object of renewed controversy — consists of neglecting the effects of magnetic diffusion in the core so that the secular variation is considered as being due entirely to the advection of magnetic field lines by the fluid motions in the core. The application of the frozen-flux approximation to reversals — a process with rather longer time-scale than is normally believed appropriate for the approximation — was first considered by Gubbins and Roberts⁸, who presented simple examples of frozen-flux reversals, and showed that axisymmetric, frozen-flux reversals are impossible. Valet *et al.* refer to the recent extension to the frozen-flux theory of the secular variation made by LeMouél⁹, who showed that the force balance at the top of the core is likely to be geostrophic. This imposes strong constraints on the flow at the core-mantle boundary and, in particular cross-equatorial flow is not permitted, a condition which makes geostrophic frozen-flux reversals very hard, if not impossible, to effect.

LeMouél also showed that with geostrophy and frozen-flux the dipole component of the field has no effect on the secular variation: Valet *et al.*'s observation that certain aspects of the secular variation, namely the long-term variations, are preserved through a reversal tends to support LeMouél's results. Moreover, they claim that this suggests that only the internal parts of the core are involved in the mechanism of a reversal. Herein lies a problem: for not only are geostrophic frozen-flux reversals very hard to effect, but if the reversal process is effectively frozen-flux then the uppermost part of the core must be involved in the process otherwise the reversed field originating deeper in the core would never be able to permeate the perfectly conducting fluid overlying it.

Alternative explanations of the persistence of these long-term variations are available. A wide variety of types of oscillatory behaviour are possible in the

core, from relatively simple hydrodynamic waves to fully nonlinear dynamo oscillations. Hide¹⁰ considered free hydrodynamic oscillations, the period of which depend on the Earth's rotation rate and the strength of the toroidal field in the core. Of these the former is certainly constant to a very good approximation through a reversal and the latter may not necessarily change greatly. The period of Hide's fundamental mode of oscillation, with the assumption of a toroidal field strength of 100 μ T, is 2,900 years.

The study of reversals seems to be typical of other branches of geophysics: as the available observations have improved, so the complexity of the process has become apparent. However, the two papers in this issue do reinforce the picture of fluid-driven reversals, characterized by a

metastable intermediate field configuration, that has been emerging over the last few years. Perhaps now the ball is back in the court of the dynamo theorists to explain the observations, but given the highly nonlinear nature of the dynamo problem, such a prospect seems rather dim. $\square$

1. Valet, J.-P., Laj, C. & Turcholka, P. *Nature* **322**, 27 (1986).
2. Herrero-Bervera, E. & Theyer, F. *Nature* (in the press).
3. Hillhouse, J. & Cox, A. *Earth planet. Sci. Lett.* **29**, 51 (1976).
4. Hide, R. *Nature* **293**, 728 (1981).
5. Prévot, M., Mankinen, E.A., Grommé, C.S. & Coe, R. *Nature* **316**, 230 (1985).
6. Shaw, J. *Geophys. J. R. astr. Soc.* **40**, 345 (1970).
7. Hoffman, K.A. *Nature* **320**, 228 (1986).
8. Gubbins, D. & Roberts, N. *Geophys. J. R. astr. Soc.* **73**, 675 (1983).
9. LeMouél, J.-L. *Nature* **311**, 734 (1984).
10. Hide, R. *Phil. Trans. R. Soc.* **A259**, 615 (1966).

Jeremy Bloxham is at the Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts 02138, USA.

Cancer

Cell growth control mechanisms

from Tony Hunter

UNDERSTANDING the cellular basis of cancer means being able to describe the biochemistry of the regulated pathways between cell surface and nucleus that control cell growth. Two concurrent recent symposia* on the activation and inhibition of oncogenic proliferation provided much new information on the balance of regulatory influences that constrain the growth of normal cells and fail in cancer.

On the basis of a simple observation made some 15 years ago, Henry Harris inferred the existence of 'tumour suppressor genes': if cells from two highly malignant tumour lines are hybridized, the resultant hybrid is usually non-malignant. This implies that many oncogenic mutations involve the loss of function (that is, that they are recessive). More recently, evidence for such recessive mutations in human tumours has emerged from studies of hereditary predisposition to cancers such as retinoblastoma and Wilms' tumour, which are characterized by the loss of a specific region of both copies of chromosome 13 or 11, respectively. These observations imply that the missing region contains a suppressor gene, and have no doubt helped to stimulate the recent resurgence of interest in suppressor genes, which three groups have been reinvestigating at the molecular level.

E.J. Stanbridge (University of California, Irvine), for example, reported on the occasional tumorigenic segregants that arise from hybrids between normal human fibroblasts and tumorigenic HeLa cells, which are generally non-tumorigenic. A common feature of these segregants is the

loss of a single copy of chromosome 11 or 14, and, by using restriction fragment length polymorphisms, Stanbridge has been able to establish that in 5 out of 6 cases the normal fibroblast chromosome 11 has been lost. The introduction of a normal human chromosome 11 (marked with a drug-resistance gene) by mini-cell fusion into HeLa cells suppressed their tumorigenicity, implying the presence of a suppressor gene on chromosome 11. Hybrids of human keratinocytes with HeLa cells, which are of epithelial origin, are also non-tumorigenic, but in this case the tumorigenic segregants have lost chromosome 1 or 4. This suggests that cells of different embryonic lineages express different suppressor genes.

How do suppressor genes work? Experiments with transfected cells show that they can act to nullify the effects of exogenously added oncogenes. J.C. Barrett (National Institute of Environmental Health Science, Research Triangle Park), for example, has isolated tumorigenic Syrian hamster embryo (SHE) cells following co-transfection with the v-Ha-ras and v-myc genes and finds that such tumour cells nonrandomly lose one copy of chromosome 15. Fusion of these cells with normal SHE cells suppresses the tumorigenic phenotype, but the level of p21^{ras} is not decreased. (The occasional retransformant generated from such hybrids had again lost chromosome 15.) Similarly, normal hamster fibroblasts (CHEF) can suppress the phenotype of CHEF cells transformed with the EJ-Ha-ras oncogene on fusion, while expression of p21^{ras} remains undiminished (R. Sager, Dana-Farber Cancer Institute, Boston). The transformed cells characteristically lack

*Growth Factors, Tumour Promoters and Cancer Genes and Interferons as Cell Growth Inhibitors. UCLA Symposia held at Steamboat Springs, Colorado, 6–13 April 1986.

part of chromosome 3, which would be restored in the hybrids. This suggests that there is a suppressor gene on hamster chromosome 3.

One obvious area of interest is the identification of natural inhibitors of growth. It now seems that both interferon and transforming growth factor (TGF- β) may fulfil this function. Moreover, it is becoming clear that many growth factors induce the expression of a β -type interferon, presumably providing an inhibitory feedback loop to prevent runaway proliferation. Tumour necrosis factor (TNF), which paradoxically is cytotoxic to many human tumour cells while being a growth factor for normal human fibroblasts, induces interferon- β , (IFN- β) (J. Vilcek, New York University Medical Center). Anti-IFN- β antibodies increase the growth-stimulatory effect of TNF, implying that the IFN- β acts as a growth inhibitor. Platelet-derived growth factor induces delayed IFN- β expression in quiescent mouse 3T3 cells (C.J. Stiles, Dana-Farber Cancer Institute). TGF- β , which like TNF has a dual stimulatory/inhibitory nature depending on the target cell, induces IFN- β in fibroblasts for which TGF- β is mitogenic, but not in epithelial cells for which it is a growth inhibitor (H.L. Moses, Vanderbilt University). Perhaps significantly, the induced IFN- β is not an interferon which is classically involved in the antiviral response, but one with a rather poor antiviral activity, possibly because its natural function is in growth control. Haematopoietic cells stimulated with mitogens also conform to this pattern. For instance IFN- γ is induced in T cells by the experimental mitogen phytohaemagglutinin (PHA) and in macrophages by the natural mitogen CSF-1.

A major signpost to pathways of tumorigenesis is provided by the evidence that one class of oncogenes is derived from the genes encoding the receptors for growth factors with protein-tyrosine kinase activities. In this connection, the most exciting result was the report of the activating mutation in the *neu* oncogene (C.I. Bargmann, Massachusetts Institute of Technology). The product of the *neu* oncogene, which is present in a series of ethylnitrosourea (ENU)-induced rat neuroblastomas, is a protein-tyrosine kinase, closely related to the epidermal growth factor (EGF) receptor, that is presumed to be a growth factor receptor, although no ligand has yet been identified. The activating event is a point mutation converting a valine to a glutamic acid in the transmembrane domain. Strikingly, the same mutation in the *neu* gene occurs in four independent tumour lines. As with other oncogene products of this type, it is assumed that the activated *neu* protein has constitutive, rather than ligand-regulated, protein kinase activity, although this re-

Tropical changes at the sea surface

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Satellite measured sea surface temperatures ($^{\circ}\text{C}$) show that on 28th June 1983 (bottom) the eastern tropical Pacific was unusually warm (during El Niño) while the eastern tropical Atlantic was cold. A year later (top) the sea surface temperature patterns in the two regions were reversed. Papers describing and interpreting the unusual conditions in the tropical Atlantic over this period will appear in *Nature* on 17th July. (Photograph from R. Legeckis, National Earth Satellite Service and O. Brown, University of Miami).

mains to be proven. If so, this clearly suggests that the transmembrane domain is much more important in transmembrane signalling by growth factor receptors than heretofore ceded.

How exactly the activation of protein-tyrosine kinases contributes to tumorigenesis has not been established. There is, however, some evidence linking their activities to the phosphatidylinositol (PI) cycle system that is known to be activated by mitogenic signals. These signals are believed to stimulate a GTP-dependent phospholipase C in the cell membrane,

leading in turn to the production of inositol trisphosphate and diacylglycerol, which activates protein kinase C, an effector on this mitogenic pathway.

It now seems there may be several ways for transforming proteins to accelerate PI turnover and the consequent activation of protein kinase C and Ca^{2+} release. C.J. Sherr (St Jude's Children's Hospital, Memphis), using a new *in vitro* assay for the membrane-associated phospholipase C, finds a GTP-dependent increase in activity in mink cells transformed by either *v-fes* or *v-fms*, both of which encode pro-

tein-tyrosine kinases. This implies that either the phospholipase itself or one of its regulatory factors may be phosphorylated and stimulated by the protein-tyrosine kinase in question. Cells transformed by *v-ras*, which does not encode a protein-tyrosine kinase, also display increased phospholipase C activity, presumably by a different mechanism. By contrast, polyoma virus transformation may stimulate one of the PI kinases, possibly through direct phosphorylation by the activated form of pp60^{src} found in association with the viral middle-T antigen (L.C. Cantley, Tufts University). This lipid kinase activity has been partially purified, which should allow the necessary biochemistry to be done.

The effector end of the cycle was addressed by two groups who reported the isolation of complementary DNA clones for protein kinase C using brain RNA as a source of messenger RNA. Evidence for a second closely related protein, which was differentially expressed, emerged from the cloning, suggesting that protein kinase C is a family of enzymes rather than a single entity. This has obvious implications for understanding the regulation of protein kinase C by its natural regulators, Ca^{2+} and diacylglycerol, and by tumour promoters acting through the diacylglycerol binding site. The predicted structure of the enzyme is largely as anticipated, with the protein kinase domain near its carboxy-terminal end, but provides no clues to the nature and location of the Ca^{2+} , phospholipid and tumour promoter/diacylglycerol binding sites. These will have to be defined by molecular manipulations of the cDNA clones.

To complete the circle, there is now evidence suggesting that some of the substrates of protein-tyrosine kinases are inhibitors of phospholipid turnover. Of two purified phospholipase A₂ inhibitors from human placenta (Barbara Wallner, Biogen Research Corporation), one seems to be identical to the known phospholipase A₂ inhibitor lipocortin, which in turn is probably the same as the EGF receptor protein-tyrosine kinase substrate p35—at least to judge from a comparison of the lipocortin cDNA with the partial amino-terminal sequence of p35 (S. Cohen, Vanderbilt University). Moreover, the sequence of the second placental inhibitor (Wallner) reveals it to be the major retroviral protein-tyrosine kinase substrate p36 (my own work and R.L. Erikson, Harvard University); and there proves to be more than 50 per cent amino-acid sequence identity between lipocortin/p35 and p36, suggesting that there is a family of such proteins.

The ultimate target of most plasma membrane signalling systems is the nucleus. The *myc* and *fos* genes, which have been shown to be rapidly activated in quiescent fibroblasts by various mitogenic

stimuli, received the most attention. The mechanism of induction of the *c-fos* gene in PC12 cells (a pheochromocytoma cell line able to differentiate into cells with neuronal properties in response to nerve growth factor, NGF) has been studied in detail by T. Curran (Roche Institute of Molecular Biology) and E. Ziff (New York University Medical Center). There seem to be at least two induction pathways: first, NGF and fibroblast growth factor induce expression of the *c-fos* gene, possibly in part through the activation of protein kinase C; and second, opening of a voltage-dependent Ca^{2+} channel and an increase in intracellular Ca^{2+} may activate a calmodulin-dependent enzyme.

Except for their localization to the nucleus, there has been a paucity of clues as to the functions of the *myc* and *fos* proteins. The independent observations by R. Watt (Smith Kline and French Laboratories) and J.M. Bishop (University of California, San Francisco), which show that *c-myc* proteins co-localize with snRPs in the nucleus by immunofluorescence staining, is intriguing and suggests that the *c-myc* protein may have a role in RNA processing. The *c-myc* and *c-myb* proteins have previously been found to bind to DNA, albeit without displaying any sequence specificity. A fraction of the *fos* protein from induced PC12 cells will bind DNA (Curran), and the *N-myc* gene product (a phosphoprotein of relative molecular mass 65,000–68,000) can do the same thing (Bishop). Perhaps the DNA binding activity of all these proteins means that they can also interact with RNA and are involved in RNA processing.

Attempts to analyse the function of the

nuclear oncoproteins by genetic means has been frustrated by the apparent absence of these genes from genetically tractable organisms such as yeast and *Drosophila*. The characterization of a putative *N-myc* gene in *Drosophila* by M.E. Lambert (Cold Spring Harbor) may make it possible to carry out the critical experiments to test its function.

The work of A. Balmain (reported on page 78 of this issue) provides a nice example of how one can do molecular biology with the whole animal. Analysis of *ras* gene mutations occurring in papillomas and squamous cell carcinomas, induced by treatment of mouse skin with an ordered combination of an initiating carcinogen and a tumour promoter, shows that the different types of point mutation in the oncogenically activated *c-Ha-ras* genes detected by transfection correlate well with those known to be caused by the initiating carcinogen. In this model of the initiation, promotion and progression events in carcinogenesis, mutation of the *c-Ha-ras* gene appears to be an early, initiating event, and can be mimicked by infection with Harvey murine sarcoma virus or by application of *v-Ha-ras* DNA to the skin followed by TPA (a caution to those who work with cloned oncogenes and TPA without gloves). Progression from a papilloma to a full-blown carcinoma is commonly associated with amplification or homozygotization of the mutant *ras* allele and presumably other events, such as loss of suppressor genes. □

Tony Hunter is in the Molecular Biology and Virology Laboratory at the Salk Institute, La Jolla, California 92037, USA.

Biochemistry

Clues about RNA enzymes

from H.D. Robertson

THREE recent papers in *Nature*, one published in May¹ and two published on pages 83 and 86 of this issue^{2,3}, present a preview of coming attractions in the study of multiple structural requirements for RNA catalysis. One emerging theme is the dynamic nature of potential RNA active sites, in which the same bases may need to undergo several kinds of secondary and tertiary structural interactions during the reaction.

The ribosomal RNA precursor from *Tetrahymena* has been shown to be capable of self-catalysed cleavage by Cech and his colleagues (ref. 4; see the *News and Views* article⁵ by F. Westheimer). Predictions by Davies *et al.*⁶ suggested a role for certain elements of secondary structure in the processing of the *Tetrahymena* rRNA precursor and other RNAs with Group I introns.

In their new work, Waring *et al.*¹ point to the importance of a short helix involving the 5' (upstream) splice site of the *Tetrahymena* rRNA precursor. They used a series of carefully selected mutants in two non-contiguous 9-nucleotide-long regions of the *Tetrahymena* rRNA precursors: one region includes the 5' splice site; the other, the internal guide sequence (IGS), is part of the intron. These data support the idea that nine Watson-Crick base pairs (bp) form between these two segments to produce a helix which spans the 5' splice site. Waring *et al.* report that single or double mutations which interrupt the pairing of this stem block cleavage at the 5' splice site. Compensatory mutations, designed to restore base pairing, lead to cleavage at the correct site. However, in one striking example, a simple transversion of two consecutive

base pairs from one strand to the other reduced the rate of RNA cleavage nearly 100-fold. The authors conclude that the RNA-RNA helix is necessary, but not sufficient, for effective 5' splicing and comment that it is very unlikely that the energy differences between two such closely related 9-bp stems could account for the observed rate differences. Instead, they suggest that "the sequence of (these segments) may play a role in positioning the helix via tertiary interactions between exposed groups on bases in (this) and other parts of the RNA catalyst".

The detailed crystal structures of various model DNA helices support this idea and indicate future directions for experimentation. In addition, the exact role of the 9-bp Watson-Crick secondary structure needs to be investigated in more detail, using techniques similar to those of Waring *et al.* to generate mutations which give aberrant cleavage, rather than no cleavage. Although a helix may help to define the splice site, precision may be conferred by other structural features, with the base pairs needed for stability.

The two papers in this issue address additional aspects of 5' splice-site specifications in Group I introns. One of these³, by Garriga, Lambowitz, Inoue and Cech, takes advantage of a previous finding⁷ that certain RNA-catalysed reactions involving the 5' splice site of the *Tetrahymena* rRNA precursor can occur in *trans*. For example, when the dinucleotide CU is incubated with the rRNA precursor in the absence of a guanosine cofactor, normal cleavage at the 5' splice site is not observed; however, the dinucleotide becomes attached to the 3' (downstream) exon at the exact 3' splice site. From this work and the studies of Waring *et al.*¹, it appears that the 5' exon remains bound to the IGS even after cleavage at the 5' splice site and that the IGS then positions the 3' splice site for cleavage. Thus, the interaction which forms the 9-bp stem may normally hold the 5' exon in position to attack the 3' splice site, but when cleavage at the 5' splice site (...CU₂) is blocked, the dinucleotide CU temporarily disrupts the structure around the 5' splice site and cleaves, adding itself to the 3' exon. This explanation for the effect of CU has been tested by Garriga *et al.*³.

The authors took a second Group I intron, which has a different sequence in the helix spanning the 5' splice site. In agreement with earlier predictions of Davies *et al.*⁶, the sequence in the *Neurospora* mitochondrial cytochrome *b* intron 1 messenger RNA (CUGGGU, instead of the CUCUCU of *Tetrahymena*) has a complementary region in a segment analogous to the IGS of *Tetrahymena*. Garriga *et al.* show that CU promotes cleavage of the *Neurospora* mRNA precursor to give a 3' exon with GU attached to its 5' end. CU is inactive in this assay. Interestingly,

Galactic strings are stable

Loops of cosmic string may be the seeds of galaxies, as many astrophysicists have suggested, but are they stable and do they last long enough to set galactic formation going? Two theorists from Imperial College, London, have just addressed this key issue, and found that the answer is yes (Copeland, E.J. & Turok, *N. Phys. Letts. B* 173, 129; 1986).

Initially the strings emerge from the mesh of vortices presumed present after the "phase transition" or "condensation" that takes place when the universe cools through the grand unification temperature — a temperature above which the strong, weak and electromagnetic forces behave as if they had equal strength. The galactic loops are chopped off the mesh about 10⁸ seconds (3 years) after the Big Bang. But the seeding of galactic matter cannot take place before the Universe passes the state of equal matter and radiation density some 10¹¹ seconds (3,000 years) later.

Thus the cosmic loops of self-gravitating string, weighing 1,000 tonnes per Fermi (10⁻¹³ cm), must last 3,000 years — or the stringy explanation of galactic formation will not work.

Basically, Copeland and Turok set out to show that a complete "seed" loop, which should weigh in at about one-hundredth of a galactic mass and be 30 parsecs long at the seeding time, will not intersect itself on a 3,000-year timescale. Other studies have shown that self-intersecting loops cut themselves in two to form smaller loops,

and if this process continued over a time-scale short compared with 3,000 years there would be no loops left of the right size to seed galaxies.

In fact, the spectrum of loops that Copeland and Turok start with at 10⁸ seconds has already ironed out simple twisting (as when an elastic band is twisted into a figure eight); the loops they begin with are those left after all initial motions that lead to self intersection have done their work. For the problem in hand, it is necessary to consider perturbations of the remaining strings — caused by gravitational effects — reaching sufficiently large amplitudes to cause self-intersection. In principle, cusps can form, for example, that will bite off small loops. Copeland and Turok's calculations show that these effects are negligible over a 3,000-year period, however, so the galactic loops remain effectively stable.

Do the loops remain in present galaxies? No. They steadily shrink, losing mass as gravitational waves, until they vanish in a puff of super-heavy bosons. Looking out into the Universe from our present vantage point, this process was occurring at a redshift of around 100. The furthest known objects are something over a redshift of 3. However, the larger and so longer-lived loops that are postulated to cause the formation of galactic clusters should still be around. They should be hiding in the clusters with lengths of a few kiloparsecs, and masses of the order of that of a galaxy. Robert Walgate

adding trinucleotides to these reactions gives less clear results, leading the authors to point out that "...although Watson-Crick base pairing with the intron binding sites may be an important determinant of the *trans*-splicing reaction, additional interactions must also contribute to determining the extent of reaction with trinucleotides. Base stacking or non-Watson-Crick hydrogen bonding might allow certain trinucleotides to interact with the normal binding site. Alternatively, the trinucleotide might interact with other, at present unidentified binding sites within the intron". Again we have here base-pairing rules as part of the explanation but additional structural elements are invoked.

In the other paper in this issue² J.W. Szostak synthesized *in vitro* a 'core' of the rRNA intron containing about 250 internal bases but neither of the splice junctions nor either of the two 9-nucleotide RNA segments whose base pairing I have described here as an important element in cleavage at the 5' splice site. Both these segments are present in a second 'substrate' molecule, a 119-base-long RNA which contains the last 32 nucleotides of the 5' exon and the first 62 nucleotides of the intron, and thus spans the 5' splice

site. Neither core enzyme nor substrate RNA is capable of self-cleavage, but gel electrophoresis shows that the core cleaves the substrate accurately to give the expected 5' end of the intron sequence. The interactions between the two molecules studied by Szostak cannot be stabilized by Watson-Crick structural interactions involving the IGS.

Furthermore, Szostak finds optimal cleavage under conditions quite different from those originally characterized by Cech and colleagues⁷ for the self-catalysed splicing of the intact *Tetrahymena* rRNA precursor. Remarkably, the optimal temperature is 58°C. Szostak concludes "it is likely that the structure of the enzyme is stabilized by numerous tertiary interactions, and that the enzyme in turn stabilizes the structure of the substrate." □

1. Waring, R.B., Townner, P., Minter, S.J. & Davies, R.W. *Nature* **321**, 133 (1986).
2. Szostak, J.W. *Nature* **322**, 83 (1986).
3. Garriga, G., Lambowitz, A.M., Inoue, T. & Cech, T.R. *Nature* **322**, 86 (1986).
4. Kruger, K. *et al.* *Cell* **31**, 147 (1982).
5. Westheimer, F. *Nature News and Views* **319**, 534 (1986).
6. Davies, R.W. *et al.* *Nature* **300**, 719 (1982).
7. Inoue, T. *et al.* *Cell* **43**, 431 (1985).

H.D. Robertson is in the Laboratory of Genetics at Rockefeller University, 1230 York Avenue, New York 10021, USA.

Geophysics

A paradigm shift in glaciology?

from G.S. Boulton

Two papers on pages 54 and 57 of this issue^{1,2} suggest that part of the west Antarctic ice sheet is underlain by a thin layer of deforming water-soaked sediment. In such areas, the rheological properties of the sediment must largely control the dynamic behaviour of the ice sheet. Until now, most models of glacier flow have assumed that the ice moves over a rigid surface, although this was clearly not true for the great mid-latitude ice sheets of the Quaternary era which flowed over strongly deformed and predominantly soft sediment beds. Perhaps we must now revise our view of the sub-glacial processes which control modern ice-sheet flow, and use the new geophysical data that have been gathered from Antarctica to improve our reconstructions of the Quaternary ice sheets.

The great naturalists of the eighteenth and nineteenth centuries established the fact of glacier movement and observed both internal flow and basal sliding. They were impressed by the erosive power of the debris-studded soles of glaciers, which ground over bedrock to produce smoothed and striated surfaces. Similar surfaces in the mid-latitude lowlands of Europe and North America were thought to support the glacial theory, which asserted that in the recent past these areas had been covered by glaciers of continental scale, similar to the modern ice sheets of Greenland and Antarctica.

Subsequent discussions about glacial-bed processes during the early years of this century concentrated on the erosion of rock beds in alpine and highland regions, as the only direct observations of the glacier/bed interface were in natural cavities on the lee-side of bedrock protuberances³. Soft sediment-floored glaciers do not permit such observations as sediment is readily squeezed into incipient subglacial cavities. Thus, it is not surprising that sliding over a rigid bed provided the first quantitative analyses of the processes of basal *decollement* of glaciers^{4,5}, and that glacier movement was generally viewed as consisting of only two possible components, internal flow and basal sliding. As late as 1979, the symposium of the Glaciological Society on glacier beds was subtitled 'the Ice-Rock Interface', as if the two were synonymous.

Such a paradigm should have been quite implausible to geologists; that it survived so long is an indication of how little note the glaciological community took of the views of geologists. A representative measure of the relative importance of different basal processes has until now been acces-

sible only by the study of the exposed beds of the last mid-latitude ice sheets in Europe and North America (where more than 80 per cent of the surface is overlain by soft deformable sediments) and in the areas exposed by recent retreat of valley glaciers, where sediment beds also predominate. A model in which a glacier slides over a smooth, passive, rigid surface seems quite inappropriate for most of these areas. The existence of widespread glacially induced deformation structures in these sediments has long been known⁶; such structures are now being more widely recognized, and are often associated with drumlins⁷ (streamlined sediment hills produced beneath glaciers). Recent experiments demonstrate that even relatively coarse-grained subglacial sediments deform readily beneath a glacier and may play a dominant role in glacier movement⁸. It has been further argued⁹ that such a mechanism controlled the behaviour of mid-latitude Quaternary ice sheets, with a soft, easily deforming sedimentary substratum permitting high glacier velocities for relatively low driving stresses; rapid responses to changing climate; and fast volumetric growth and decay facilitated by rapid changes in the extent of the ice sheets.

The modern ice sheets of Antarctica and Greenland are thought to be quite different: large, sluggish masses moving by internal flow with fast ice streams radiating through them¹⁰ and discharge (in Antarctica) up to 90 per cent of the mass flux, although they comprise only 13 per cent of the ice-sheet perimeter. It is assumed that sliding over bedrock occurs beneath these ice streams with a lubricating water layer that decouples the glacier from its bed, thereby reducing friction and permitting fast, low-stress sliding^{11,12}. A predominantly rock bed could result from progressive stripping of pre-existing sediment during the 15 million years of continuous ice-sheet residence on the Antarctic continent.

The recent exciting results from Antarctica reported in this issue^{1,2} suggest that this picture may not be correct. The results are particularly important in demonstrating the relationship between a possibly drumlinized subglacial sediment stratum and a fast ice stream, and illustrate a geophysical technique that could establish the distribution of their physical properties and their relationship to ice-sheet dynamics on a wide scale.

Two important implications should be explored: first, that Pleistocene drumlin fields reflect widespread deformation of

soft subglacial sediments, offer little frictional resistance to ice movement and are therefore geological reflections of former fast ice streams; and, second, that soft subglacial sediments are important in regulating the response of the modern Antarctic ice sheet to climate and sea-level changes, a matter of concern even in the short-term future¹³. They may also be important in glacier-surging behaviour.

To explore the implications of subglacial sediment deformation for ice-sheet behaviour we need to understand more about the rapid but sustained deformation of sediments at very low effective stresses, in which dilation is important. Sadly, although such processes are widespread in gravitational flows of many types, adequate flow laws have not been developed and the small strains in most conventional geotechnical experiments make these experiments an inappropriate basis for such laws. Strain rates are strongly affected by water content, so the behaviour of the interstitial water system is of critical importance.

The subglacial zone is different from the *decollement* zone in most other gravity-flow systems, such as gliding nappes and mud flows, in that the glacier sole is a water source, the glacier is an acquiclude and the subglacial water escape pathways are long. Thus, high pore-water pressures need not be transient but can be sustained in a steady state.

The fundamental conceptual difference between analyses of fast low-stress glacier movement over rigid or deformable surfaces is that whereas in the former the surface can be regarded as passive (although it may be temporarily changed by the thickening of a subglacial water film), the strong interaction in the latter requires a coupled analysis of both ice and sediment deformation. Moreover, in coupling the form and structure of the bed with the dynamics of the glacier, such analyses might help to unify glaciology and glacial geology, to the benefit of both disciplines, in a way which until now has been sadly lacking. □

1. Blankenship, D.D., Rooney, S.T., Bentley, C.R. & Alley, R.B. *Nature* **322**, 54 (1986).
2. Alley, R.B., Blankenship, D.D., Bentley, S.T. & Rooney, S.T. *Nature* **322**, 57 (1986).
3. Carol, H. *J. Glaciol.* **2**, 57 (1947).
4. Weertman, J. *J. Glaciol.* **21**, 33 (1957).
5. Liboutry, L. *Ann. Geophys.* **15**, 250 (1959).
6. Slater, G. *Proc. Geol. Ass.* **37**, 392 (1926).
7. Stanford, S.D. & Mickelson, D.M. *J. Glaciol.* **109**, 220 (1985).
8. Boulton, G.S. & Jones, A.S. *J. Glaciol.* **90**, 29 (1979).
9. Boulton, G.S., Smith, G.D., Jones, A.S. & Newsome, J.J. *geol. Soc. Lond.* **142**, 447 (1985).
10. Hughes, T.J. *Rev. geophys. Space Phys.* **78**, 1 (1977).
11. Rose, K.E. *J. Glaciol.* **90**, 63 (1979).
12. McIntyre, N.F. *J. Glaciol.* **108**, 99 (1985).
13. *Glaciers, Ice Sheets and Sea Level: Effects of a CO₂ Induced Climatic Change* (US Department of Energy, Washington DC, 1985).

G.S. Boulton is a Reader in Environmental Sciences at the University of East Anglia, Norwich, UK, and Professor of Physical Geology at the University of Amsterdam, Spui 21, 1012 WX Amsterdam, The Netherlands.

A new term in the global carbon balance

SIR—The yearly carbon balance is unbalanced: man-made emissions of carbon dioxide into atmosphere, through combustion of fossil fuels and deforestation, are not fully accounted for by ocean adsorption and measured increase of CO_2 in the atmosphere. Man can operate powerful disturbances on ecosystems; we will show he can also operate powerful segregation of carbon through a multiplicity of actions which altogether we propose to call "stock and waste" segregation.

Confining ourselves to the most recent estimates of anthropogenic emissions of carbon into the atmosphere, we have a flow of 5.2 gigatons (10^{15} g) carbon per year coming from fossil fuel combustion¹ and a flow from deforestation estimated either to be $0.9\text{--}2.5 \times 10^{15}$ g C yr^{-1} (ref. 2) or $1.6 \pm 0.8 \times 10^{15}$ g C yr^{-1} (ref. 3). The classical sinks for this total C flow, ocean adsorption (maximum rate 2.5×10^{15} g C yr^{-1}) and increase of CO_2 in the atmosphere (measured rate of 2.3×10^{15} g C yr^{-1}) are evidently not large enough to accommodate for emissions. We therefore need one or more new carbon sinks.

While revision of older mechanisms is going on and new mechanisms await better understanding and quantification³, we wish to draw attention to a new carbon segregation path whose existence seems to us beyond any doubt.

The fundamental starting point is to recognize that each of the 4,000 million people in the world, within their social and economic structure, is entitled to the possession or use of an array of commodities and services, each having a material base and therefore a carbon content. The ensemble of these possessions, which in developed countries covers an impressive variety of objects, is what we call stock.

In societies recording a continued increase in the standard of living after correction for population growth, stock increases every year.

The carbon content of the global annual increase of stock is a sink that needs evaluation. But commodities and services do not last for ever. Goods must be renewed and new products displace older ones. This creates an enormous flow of materials — the flow of waste.

Whenever waste is not somehow "recycled" but, as in landfilling or dumping, essentially subtracted from natural cycles, its carbon content is segregated for many years. This represents the waste component of our stock and waste segregation mechanism.

Our group is still working on the quantitative evaluation of the stock and waste mechanism in terms of the amounts of carbon that are segregated every year. We

can, however, give some preliminary results which show that stock and waste must be taken into account as a term in a correct carbon balance of the world.

As an example of stock increase, we took into consideration the Food and Agricultural Organization⁴ data on wood production in 1982, during which the wood and pulp and paper industries dealt with 0.7×10^{15} g of carbon. Looking at final uses of this carbon⁵, at least 0.45×10^{15} g C yr^{-1} are accumulating in the stock, whereas the rest enters the flow of waste.

Considering the waste contribution to the carbon sink, we note that despite the fact that we are the only producers, our knowledge of global waste production is very uncertain.

Recent data on 10 European countries with a population of 270 million people show that total waste reaches a value of 2 gigatons per year (W. Ganapini, personal communication); a very conservative extrapolation gives a value of at least 7 gigatons for the world.

Heating values for wastes are in the range 800–3,000 kcal kg^{-1} (ref. 6); an average value of 1,200 kcal kg^{-1} gives, according to the correlation shown by Tillman⁵, a content of 10% of carbon; assuming that 50% of this waste is subtracted from any substantial oxidation, waste is responsible for the segregation of an impressive amount of carbon, at least 0.35×10^{15} g C yr^{-1} .

Thus the evaluation of just a few terms of the stock and waste mechanism already shows its importance: anthropogenic contribution to carbon subtraction ranks of some 10^{15} g every year. Carbon balance must be accordingly modified.

UMBERTO BIANCHI

LUIGI MUSI

MARINA PICARDI

EUGENIO PIOVANO

*Institute of Industrial Chemistry,
University of Genoa,
C.so Europa 30,
16132 Genoa, Italy*

1. Woodwell, G.M. *et al. Science* **222**, 1081 (1983).
2. Houghton, R.A. *et al. Nature* **316**, 617 (1985).
3. Bolin, B. *The Greenhouse Effect, Climatic Change and Ecosystems* (Wiley, Chichester, in the press).
4. Production Yearbook, FAO (Rome) 1982.
5. Tillman, D.A. *Wood as an Energy Resource* (Academic, London, 1978).
6. Cassitto, L. & Magnani, P. *Ambiente e Sicurezza*, 44 (November 1983).

X chromosomes and dosage compensation

SIR—The contention of Chandra^{1,2} that the primary function of X-chromosome inactivation is one of sex determination rather than dosage compensation is difficult to justify for more reasons than those considered by Lyon³. When examined in the context of the homomorphic sex chromosomes of diploid dioecy and in relation

to the X and Y of haploid dioecy, the thesis is untenable.

Sex determination which depended solely on differential inactivation in one or other of a pair of homologues must be recognized as environmental or phenotypic. There is therefore no reason to suppose that alleles related to secondary sex-specific characters would accumulate preferentially in either homologue, for each chromosome could play the role of an X or Y in different individuals. Genetic fixation of sex determination is thus an essential precursor to the evolution of heteromorphic sex chromosomes in an XX/XY or ZZ/ZW system, and the same is true of the haploid X/Y system of bryophytes^{4,5}.

In many mosses and liverworts there are reports of homomorphic sex chromosomes where the Y of male plants differs from the X of females only in the greater extent of its heterochromatin which, there is evidence to suggest⁶, is facultative. It has been seen in terms of more extensive gene inactivation of secondary female characters in the Y than of male characters borne by the X, characters which it is hypothesized are thus located as a consequence of an earlier monoecious condition^{4,5}. If sex-determining factors were not chromosomally fixed, both of these homologues would be variously phenotypic X or Y chromosomes. Evolution of heteromorphic X and Y chromosomes, as seen in many families and genera, would not then have followed. Instead, limited but equal erosion of the X and Y would be an expected consequence if crossing-over were excluded from lengths of chromosomes bearing secondary sexual characters⁷, whereas no significant erosion would occur in the absence of chiasma localization.

M.E. NEWTON

*Department of Botany,
University of Manchester,
Manchester M13 9PL, UK*

1. Chandra, H.S. *Proc. natn. Acad. Sci. U.S.A.* **82**, 6947–6949 (1985).
2. Chandra, H.S. *Nature* **319**, 18 (1986).
3. Lyon, M.F. *Nature* **320**, 313 (1986).
4. Newton, M.E. in *The Experimental Biology of Bryophytes* (eds Dyer, A.F. & Duckett, J.G.) 65–96 (Academic, London, 1984).
5. Newton, M.E. *Bull. Br. bryol. Soc.* **45**, 9 (1985).
6. Newton, M.E. in *New Manual of Bryology Vol. 1* (ed. Schuster, R.M.) 117–148 (Hattori Botanical Laboratory, Nichinan-shi, 1983).
7. Bull, J.J. *A. Nat.* **112**, 245–250 (1978).

Motions of the images of atoms

SIR—Howie might have missed an alternative explanation for motions of atom images under the electron microscope in his article "Coulomb explosions in metals" (*Nature* **320**, 684; 1986). The observations made by the groups whose work he reports consist of apparent movement of atoms or columns in gold microcrystals when imaged in the electron

microscope using the rather high electron flux rates required for high resolution. Undoubtedly, there must exist fiendishly high electric field gradients on a highly local level caused by electron-atom collision-induced ionization, and they might be responsible for some positional instabilities of the atoms.

However, one should also consider the effect these charges in the specimen might have on the trajectories of subsequent incoming electrons. They would most certainly distort the optics on a local level. Furthermore, as the charges build up and dissipate — by whatever mechanism(s) — the fluctuating local electric fields cause fluctuating local distortions in the microscope's optics (A.P.K. *Ultramicroscopy* 7, 351–370; 1982). The net result would be somewhat like looking at a photograph lying (face up) on the bottom of a pool of water whose surface is inundated by random ripples. Is it the painting that is undergoing 'convulsions', or is it just its image?

ALEX P. KORN

Biochemistry Department,
Weizmann Institute of Science,
76100 Rehovot, Israel

HOWIE REPLIES—Specimen charging effects certainly can disrupt the imaging process in the electron microscope, giving rise for example to sudden and large image displacements. It seems unlikely, however, that this can account for the majority of the observed apparent changes of structure, particularly changes from single crystal to multiple twin and back again.

A. HOWIE

Department of Physics,
Cavendish Laboratory,
Madingley Road,
Cambridge CB3 0HE, UK

Legal problems of Huntington's chorea tests

SIR—Recent correspondence on the ethics of using a recombinant DNA probe to diagnose Huntington's chorea has not mentioned counsellors' potential legal obligation to provide such a test, regardless of their personal ethical judgement. That the G8 probe does not provide accurate diagnosis is irrelevant here, as all medical diagnoses provide only evidence for disease states, not proof of them. An established genetic screen, fetal karyotyping to detect Down's syndrome, can give good evidence that a fetus is normal, but can also give false negatives for a mosaic fetus¹ or for one carrying a Down's-causing translocation², as well as through sampling or laboratory errors. Counselling about the test for at-risk mothers includes mention of these risks, as will analogous counselling for the G8 probe³.

The expectation that mothers over 35 years old will receive fetal karyotyping, with attendant counselling, has recently

acquired a legal standing, with the award of £35,000 damages to Mrs Ayten Yagiz, who sued the City and Hackney Health Authority after giving birth to a Down's syndrome daughter when she claimed not to have been offered fetal karyotyping⁴. This would appear to establish a precedent for saying that doctors have a legal liability to provide established genetic tests for at-risk mothers. At 40 years old, Mrs Yagiz daughter's risk of Down's was $\approx 1\%$ (ref. 5), far less than that facing most parents seeking counselling for Huntington's chorea, cystic fibrosis or other inherited diseases for which cloned probes are becoming available. Other analogous cases are currently awaiting trial. Once cloned probes join fetal karyotyping as an established genetic screen, genetic counsellors may find they have an obligation to use them beyond their present research facilities' ability to do so. This may be a constraint on putting such probes into clinical use.

WILLIAM BAINS

Department of Biochemistry
University of Bath,
Claverton Down,
Bath BA2 7AY, UK

1. Stern, C. *Genetic Mosaics and other Essays*, 68–69 (Harvard University Press, Boston, 1969).
2. Fraser Roberts, J.A. *An Introduction to Medical Genetics*, 182–187 (Oxford University Press, 1973).
3. Watt, D.C., Lindenbaum, R.H., Jonasson, J.A. & Edwards, J.H. *et al. Nature* 320, 21–22 (1986).
4. Illman, J. *Daily Mail* 7 November 1985.
5. Cavalli-Sforza, L.L. & Bodmer, W.F. *The Genetics of Human Populations*, 101 (Freeman, London, 1971).

Is cannibalism all in the mind?

SIR—Behrensmeier *et al.* show that trampling of bones in sand can produce marks mimicking cutmarks made by flake tools. They comment that the evidence for meat-eating among early hominids is called into question by this finding.

Putative cutmarks on human bones found at various historical sites, notably at Knossos², have been used to advance the hypothesis that cannibalism was practised in these cultures. The finding of Behrensmeier *et al.* raises the possibility that this hypothesis, too, is unfounded and strengthens Arens³ contention that apparent evidence for ritual human cannibalism is in fact a psychological phenomenon of anthropologists rather than a dietary phenomenon of 'barbaric' civilizations.

R.J. BIRD

School of Behavioural Science,
Newcastle-upon-Tyne Polytechnic,
Northumberland Road,
Newcastle-upon-Tyne NE1 8ST, UK

1. Behrensmeier, A.K., Gordon, K.D. & Yanagi, G.T. *Nature* 319, 768–771 (1986).
2. Warren P. in *Proceedings of the Swedish Institute in Athens*, Stockholm (eds Hagg, R. & Marinatos, N.) (1981).
3. Arens, W. *The Man-eating Myth: Anthropology and Anthropophagy* (Oxford University Press, New York 1973).

Linkage between the nation states

SIR—The 'self-thinning' rule of plant ecology^{1,2} states that as communities of plants mature, increases in the mean mass per individual plant (W) tend to be accompanied by decreases in the number of plants per unit area (N), and in particular that

$$\log W = A + B \log N$$

where A and B are constants. If B took the value -1 then decreases in N due to self-thinning would be exactly matched by increases in W and the mass of standing crop per unit area ($=W.N$) would be constant ($=A$). In practice^{3,4}, B tends to take values close to $-3/2$, so that

$$\log W = A - 3/2 \log N$$

Hence

$$\log W = 3 \log(W.N) - 2A$$

and the standing crop increases with W . Denness⁵ suggests that the ' $-3/2$ rule' applies to the distribution of size and density between the states of Europe; and hence he argues that if the total population is to expand the number of nation states in a given area must decrease. The evidence rests on a log-log graph in which W was equated with nation size (population) and N with $1/(\text{national area})$ (that is, with number of nations per unit area)⁶. The graph resembled a scatter plot, but the addition of a grid of lines of slope $-3/2$ yielded some apparently meaningful associations between nations⁶. I suggest that any resemblance with the ' $-3/2$ rule' of plant ecology is fortuitous.

The figures for population density of

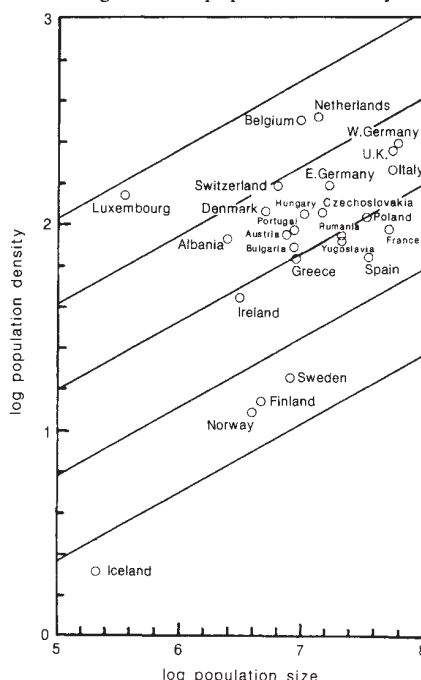


Fig. 1 A plot of $\log_{10}$ population density (people per km^2) against $\log_{10}$ population size (people per nation). The added lines have a slope $1/3$.

different nations conceal substantial fluctuations; in cities the number of people per square kilometre may be greater than in rural areas by several orders of magnitude. It seems reasonable to suppose that bigger nations will support bigger conurbations than small, and that hence the mean population density for a large nation may be greater than for smaller ones. If this were so, then nations of similar character might fall into alignment on a graph of log population density on log nation size (Fig. 1). They would also give apparently meaningful alignments in graphs of log nation size (W) on log area ($=1/N$) (ref. 6). In particular, if the slope of the graph of log population density ($=W.N$) on log nation size ($=W$) is approximately $1/3$, then the slope of the graph of log nation size (W) on log area ($=1/N$) will be approximately $3/2$; but this does not mean that people obey the same 'laws' of nature as plants. The alignments, I suggest, are attributable to the factors that control the aggregation of people in cities and not to any inherent tendency for the nation states within a continent to undergo 'self-thinning'.

RICHARD C. HARDWICK
National Vegetable Research Station,
Wellesbourne,
Warwick CV35 9EF, UK

1. Tadaki, Y. & Shidei, T. *Nippon Rin Gakkaishi* **41**, 341-349 (1959).
2. Yoda, K., Kira, T., Ogawa, H. & Hozumi, K. *J. Inst. Polytech. Osaka City Univ.* **14**, 107-129 (1963).
3. Gorham, E. *Nature* **279**, 148-150 (1979).
4. Westoby, M. *Adv. Ecol. Res.* **14**, 167-225 (1984).
5. Denness, B. *Nature* **320**, 491 (1986).
6. Hayton, A. F. *Nature* **310**, 178 (1984).

Is Classic Coca-Cola the real thing?

SIR—During 1985, the Coca-Cola Company discontinued production of its "old" Coca-Cola soft drink and introduced the New Coca-Cola. In response to public demand for the "old" Coke, the Coca-Cola Company began marketing the Classic Coca-Cola. Consumer complaints about difference in taste between the Classic and "old" version were supported by claims from both the Sugar Association, Inc.¹ and *Newsweek*² that the Coca-Cola Company had discontinued the use

of sucrose (table sugar) in 1984. Since the Classic Coca-Cola container bears the words "Original Formula" and the major ingredient other than carbonated water in both the Classic and the New Coca-Cola is "high fructose corn syrup and/or sucrose", while that of "old" Coca-Cola was "sugar", we determined the sugar content of the newly introduced drinks to compare them with our previous analyses of "old" Coca-Cola.

Sugar analyses, using a gas-liquid chromatographic procedure³, confirm claims^{1,2} that neither Classic nor New Coca-Cola contain sucrose, whereas in 1983 we found that "old" Coca-Cola contained 4.7% sucrose (Table 1). There was also no sucrose in what we refer to as "transition" Coke since it was produced just prior to the introduction of the New Coca-Cola on 23 April 1985. Total sugar content of "transition", Classic and "old" Coca-Cola was very similar, but the new version contains about 10% more total sugars than any of the others.

As noted above, the Coca-Cola labels now read "high fructose corn syrup and/or sucrose", thus eliminating the use of the word "sugar". Recently the cereal industry made a similar move by eliminating the word "sugar" from the name of cereals and from its advertising even though the product and the sugar content remained unchanged⁴. Why discontinue the use of the word "sugar" which most consumers understand to be table sugar? Because health conscious consumers often associate sugar with obesity, diabetes mellitus, heart disease and dental caries⁵.

Sucrose (table sugar) has long been considered the mortal enemy of the teeth by virtue of the amount ingested and its pattern of use⁶. What impact upon dental health should one expect from the replacement of sucrose with fructose and glucose in soft drinks? Such an assessment is impossible at the present time because there are no definitive studies which demonstrate clearly the contribution of sucrose, fructose, or glucose in soft drinks to incidence of dental caries.

If sucrose is the arch criminal of dental caries as has been claimed⁷, one might expect a decrease in the incidence of

dental caries due to the elimination of sucrose from the soft drinks. However, if fructose and glucose are as cariogenic as sucrose in soft drinks, one might expect an increase in the incidence of dental caries due to an increased amount of total sugar as seen in the New Coca-Cola. What effect, if any, the changes will have on consumer demand and/or health remains to be seen.

This work was supported by Veterans Administration Research Service. We thank R.W. Christensen and H.V. Nguyen for technical assistance and D.M. Larson for secretarial assistance.

LEIGH E. WYBORN
IRA L. SHANNON

Oral Disease Research Laboratory,
Veterans Administration Medical Center,
Houston, Texas 77211, USA
and
Department of Biochemistry,
University of Texas Dental Branch,
Houston, Texas 77225, USA

1. The Sugar Association, Inc. *The Houston Post*, 7B (15 August, 1985).
2. *Newsweek*, 54-55 (26 August, 1985).
3. Li, B. W. & Schuhmann, P. J. *J. Food Sci.* **45**, 138-141 (1980).
4. Liebman, B. *Nutr. Action* **12**, 10-11 (Jul/Aug 1985).
5. Shannon, I. L. *Brand Name Guide to Sugar* (Nelson-Hall, Chicago, 1977).
6. Shannon, I. L. *Amer. Chiropr.* **1**, 54-57 (1978).
7. Newburn, E. *Odont. Rev.* **18**, 373-386 (1967).

Are we all out of Africa?

SIR—Additional information on world population distributions of human marker genes, such as that provided by Wainscoat *et al.*¹ for five closely linked polymorphic restriction enzyme sites in the β -globin gene cluster, is always welcome. One must agree, on the basis of the restriction enzyme evidence provided, that there is a clear distinction between African and Eurasian populations, an opinion in accord with at least some other genetic studies. While Wainscoat *et al.* are nominally cautious about a second proposition, stating that their data are "consistent with" the notion that anatomically modern man (*Homo sapiens sapiens*) arose in Africa and subsequently spread to Eurasia and the Americas, it is difficult to deny that their results will be taken as "new evidence that the origin of modern man lies in Africa"².

A close scrutiny of this new evidence is especially important because of its potential value when the prime archaeological rationale for asserting great antiquity for South African *H. s. sapiens* at Klasies River Mouth and Border Cave is not conclusive^{3,4}, has been challenged⁵, and the *H. s. sapiens* status of the Omo specimens from Ethiopia has been questioned⁶.

On the basis of a sample perhaps more limited than the authors concede, Wainscoat *et al.* note that three haplotypes (the specific sequence of appearance [+]) or

Table 1 Sugar analysis of Coca-Cola by gas-liquid chromatography

Product (code)	Date of production	Date of analysis	Sugar content (%)			
			Fructose	Glucose	Sucrose	Total
"Old" (032483289)	24 Mar. '83	12 Apr. '83	3.3±0.1	2.9±0.3	4.7±0.3	10.9
"Transition" (no code)	Early 1985 before 23 Apr. '85	13 Sept. '85	6.3±0.1	4.5±0.1	0	10.8
New Coca-Cola (061485189)	14 June '85	13 Sept. '85	7.0±0.1	4.9±0.1	0	11.9
Classic Coca-Cola (K8H22)	22 Aug. '85	13 Sept. '85	6.3±0.1	4.6±0.1	0	10.9

Sugar contents are mean ± standard deviation of triplicate determinations.

absence [—] of the five sites along one chromosome), namely (+----), (---++) and (-++-), are common in Eurasia, while a fourth, (----+), is common in Africa and absent elsewhere. A punctuationalist evolutionary explanation is advanced: that a 'small' founder population of *H. s. sapiens* migrated from Africa to Eurasia, losing the characteristic African haplotype (----+) en route to Eurasia by genetic drift. Present-day 'new' haplotypes, that is, other than the four common haplotypes assumed to "predate the racial divergence", may in most instances "be derived from the four common ones by single crossovers".

Forty years ago Fisher¹ first applied the crossover hypothesis in human genetics to the rhesus (Rh) blood group system, suggesting less common Rh types "are maintained [our italics] by such occasional cross-overs" of the more common types in the British population. Wainscoat *et al.* propose a similar origin for haplotypes (+---++) and (-++++) in Melanesia and Polynesia, presumably from (+-----) and (-++++) for the former and (-++++) and (-++++) for the latter. Since the Eurasian recombinants' frequencies are 2.0% and 1.1%, while the putative parental frequencies are 65.6%, 19.8% and 6.7%, the explanation has distinguished precedent. But in Africa, which Wainscoat *et al.* do not discuss, using analogous logic, (-+---) should be a recombinant form of the common haplotypes (-+---) and (-++-). Curiously, this recombinant's frequency in Africa is 19.7%; the frequencies of the presumptive parental types in Africa are 9.8% and 1.6%. In fact, the recombinant's frequency in Africa is the same as the frequency of the second most common haplotype, (-+---), in Eurasia.

We are certain that Wainscoat *et al.* would agree that their data are limited and interpretations preliminary. Nevertheless, we believe that unwarranted credence accrues to the sequence of events they outline, however hedged, when it is presented simply as the loss of one haplotype in a founder population's move from Africa to Eurasia. It ignores the need to interpret the 20% frequency of an apparent recombinant form in Africa, and the ostensible reduction of all three common haplotypes in Africa: to 4.9% (+-----), 9.8% (-+---) and 1.6% (-+---) from the current Eurasian frequencies 65.6%, 19.8% and 6.7%, respectively. Jones and Rouhani² in their imaginative "Out of Africa" scenario see no need to explain how, for example, the common haplotypes (+-----), (-+---) and (-+---) went from 75% in their proposed "ancestral African populations" to their current combined frequency in Africa of 16.4%.

Our view is not that the data are incom-

patible with the proposition that *H. s. sapiens* evolved in Africa and moved to Eurasia, but that they do not provide any evidence for that, and are, in fact, equally compatible with the position that *H. s. sapiens* originated in Eurasia and migrated to Africa, or perhaps other alternatives. It could be argued that in Africa not only (-+---) but also the common haplotype (----+) are recombinants: Wainscoat *et al.* suggest that (----+) in Papua New Guinea may have arisen through recombination, even though it could not be by a single crossover between any of the three common Eurasian haplotypes. If so, one might ask whether the African population, characterized (80%) by thus possibly recombinant forms, is more likely to be ancestral or derived. We believe the genetic data provided by Wainscoat *et al.* are at the very least a toss-up in terms of African or Eurasian *H. s. sapiens* origins, and that the genetic distance analysis presented in their Fig. 1 dendrogram, an analytical approach notorious for its pitfalls³, speaks basically to what is not at issue, namely that the Eurasian and African frequencies are very different.

It should not be forgotten, as Antonarakis *et al.*⁴ note, that while the mechanism for producing nonrandom association of DNA sequences located in the region of the β -globin gene cluster is unknown, possible explanations include selection pressure for specific DNA sequences within the region. Selection and migration¹⁰ may be a better model for the observed data than reconstructions based on an assumption of drift and migration alone, with direction of migration emerging from, rather than imposed on, such a study.

EUGENE GILES

STANLEY H. AMBROSE

Department of Anthropology,
University of Illinois,
Urbana, Illinois 61801,
USA

1. Wainscoat, J.S. *et al.* *Nature* **319**, 491–493 (1986).
2. Jones, J.S. & Rouhani, S. *Nature* **319**, 449–450 (1986).
3. Rightmire, G.P. *Curr. Anthropol.* **20**, 25–35 (1979).
4. Howell, F.C. in *The Origins of Modern Humans* (eds Smith, F.H. & Spencer, F.) xiii–xxii (Liss, New York, 1984).
5. Binford, L.R. *Curr. Anthropol.* **27**, 57–62 (1986).
6. Wolpoff, M.H. *Paleoanthropology* (Knopf, New York, 1980).
7. Fisher, R.A. & Race, R.R. *Nature* **157**, 48–49 (1946).
8. Morton, N.E. & Lalouel, J.M. *Am. J. hum. Genet.* **25**, 422–432 (1973).
9. Antonarakis, S.E., Boehm, C.D., Giardina, P.J.V. & Kazanian, H.H. *Proc. natn. Acad. Sci. U.S.A.* **79**, 137–141 (1982).
10. Hiorns, R.W. & Harrison, G.A. *Man* **12**, 438–445 (1977).

WAINSCOAT *ET AL.* REPLY—We agree that data on β -globin gene haplotypes¹ should be analysed in relation to their evolutionary implications for human populations. Giles and Ambrose do not question our conclusion that there is a major division of human populations into African and Eurasian groups, but they raise the im-

portant issues as to which haplotypes and indeed which populations are ancestral. We believe that a small number of common haplotypes predated the racial divergence and that many of the rare haplotypes have since arisen by crossovers between them. Giles and Ambrose emphasize the fact that the second most common haplotype in Africa (-+---) could result from recombination between two haplotypes (-+---) and (-+---) which, although common in Eurasia, are relatively uncommon in Africa. It is, however, also possible that this African haplotype (-+---) is derived from the most common African haplotype (----+) by a single point mutation rather than by recombination. Fortunately, these questions may be answered by sequence analysis of the haplotypes as demonstrated by the elegant study of the 'R' and 'T' haplotypes in a segment of the β -globin gene cluster².

Of course, the data on the β -haplotypes taken by themselves do not prove an African origin for modern man. Nevertheless, this hypothesis is consistent not only with fossil evidence³, but also with the available molecular data. The two lines of evidence that are emerging from DNA analysis of human populations are, firstly, genetic distance analyses, based on allele and haplotype frequencies, which show a distinct African lineage^{4,5}, and second, data indicating a greater nucleotide sequence diversity at particular loci in African populations^{4,5}. Our β -haplotype data and more recent population data on an α -globin haplotypes⁶ both suggest greater diversity in African populations, making an African origin for man more likely than one in Eurasia.

These observations are preliminary and will require confirmation by analysis of other world populations and investigation of many other loci. Nevertheless, if they are corroborated, it would be very difficult to postulate that Eurasian populations were ancestral to African populations. Much more likely, as Giles and Ambrose put it, "Out of Africa".

J.S. WAINSCOAT

A.V.S. HILL

S.L. THEIN

J.B. CLEGG

Department of Haematology and
MRC Molecular Haematology Unit,
John Radcliffe Hospital,
Oxford OX3 9DU,
UK

1. Wainscoat, J.S. *et al.* *Nature* **319**, 491–493 (1986).
2. Maeda, N., Bliska, J.B. & Smithies, O. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5012–5014 (1983).
3. Stringer, C. in *Honinid Evolution* (ed. Foley, R.) 55–83 (Academic, New York, 1984).
4. Johnson, M.J., Wallace, D.C., Ferris, S.D., Rattazzi, M.C. & Cavalli-Sforza, L.L. *J. molec. Evol.* **19**, 255–271 (1983).
5. Chakravarti, A., Phillips, J.A., Mellits, K.H., Buetow, K.H. & Seeburg, P.H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6085–6089 (1984).
6. Higgs, D.R. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).

Behind the bifocals

Roy Porter

Franklin of Philadelphia. By Esmond Wright. *Harvard University Press:1986.* Pp.404. \$25, £21.25.

BENJAMIN Franklin is safely installed in one of the most prominent niches in science's pantheon. An early dabbler in experiments, the Philadelphia printer had put electrostatics on a new footing by the time he was forty. He advanced the theory that electricity constituted a single fluid, explained the conservation of charge, and introduced the concepts of plus and minus, positive and negative. But harnessing that terrifying force was no less vital to him than understanding it. Through demonstrating by experiment that lightning was electrical, Franklin was able to pioneer the lightning conductor (though George III stubbornly refused to employ the gadget of the Yankee rebel). And by explaining the charges of the Leyden Jar, he anticipated times when electricity would constructively lie at man's service (he himself tried shock treatment in cases of paralysis, but doubted the lasting efficacy of electrotherapy). As thumbnail sketches go, it is accurate enough to see Franklin as the founder of modern electrical science.

Yet he was also a scientific all-rounder with a ready gift for invention. He studied the Gulf Stream, developed bifocal glasses, rationalized spelling and fireplace design, and even improved the rocking-horse (how like his English contemporary Erasmus Darwin!). A jack-of-all-trades, it was his rare good fortune to be master of many.

But Franklin was never more than a part-time scientist. And it is the achievement of Esmond Wright's polished biography that, recognizing that Franklin's scientific career has already been fully analysed, he has concentrated on painting a rounded portrait of all the facets of this supremely versatile and talented man. Franklin as business-man, journalist, raconteur, politician, diplomat, sage, architect of American independence and man of the world: all are here. Not least, we see Franklin as author and embodiment of the American dream. His is an extraordinary tale of self-help and hard work leading to success.

Born in 1706 of poor immigrant stock, Franklin trained as a printer; publishing led to writing, and Franklin soon won himself renown as the author of *Poor Richard's Almanac*, full of wise saws for getting on ("Early to bed, early to rise..."). By mid-life he was wealthy enough to retire from business. Ever energetic, he threw himself into public

affairs. Prominent in founding the American Philosophical Society and what was to become the University of Pennsylvania, he was elected to the colony's assembly in 1746. Soon he found himself sent to England, in effect as ambassador for the Colonies, at precisely the moment when relations between them and Westminster were going sour.

The canny Franklin was all for compromise not confrontation. In any case, he was profoundly Anglophile, envisaging America's greatness lying within a glorious global British Empire. But it was not to be. Extreme counsels prevailed on both sides of the Atlantic, Parliament became intransigent and Franklin found himself forced to abandon pragmatism for patriotism. Becoming the elder statesman of the War of Independence, he ensured that France entered the war on the rebels' side, helping to tip the balance. He was sure of American victory, for there was no hold-

ing back progress. New Worlders were ambitious, no-nonsense and free from the fossilized forms of the Old. Combining realism and optimism, Franklin nailed his colours to the mast of the future.

Scientific biographers all too often forget that their subjects are flesh and blood creatures, with lives outside the laboratory. Franklin certainly was, with his eye for women, a taste for old madeira and an unquenchable thirst for life. Above all, Wright shows how science and politics were inextricably linked in Franklin's outlooks and career. Natural science was the technique of improving man's material lot; politics the art of social progress. "O that moral science were in as fair a way of improvement", lamented Franklin, "that men would cease to be wolves to one another, and that human beings would at length learn what they now improperly call humanity".

Amen. Ben Franklin was full of such maxims, no less wise for being homely. Anyone wishing to penetrate behind the bifocals, into the urbane mind of the first and greatest of the Yankee scientists, cannot do better than to sample this feast of a biography. □

Roy Porter is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BP, UK.

Only natural?

Alastair Fitter

The Oxford Dictionary of Natural History. Edited by Michael Allaby. *Oxford University Press:1986.* Pp.688. £20, \$29.95.

THE two essentials of a good dictionary are comprehensive coverage of its subject, and accurate and up-to-date definition of terms. The editor of *The Oxford Dictionary of Natural History* (who, incidentally, has only edited it, and not provided definitions) states that it takes in the "earth sciences, atmospheric sciences, genetics, cell structure and function, biochemistry, parasitology and other disciplines". Curiously missing from this list is ecology, once described as scientific natural history.

The natural history component is in fact supplied by copious and extensive references to taxa of living organisms, complete (apparently, and I found no exceptions) down to family level and with many genera and species of greater than average interest included as well. These entries make up at least half of the total, and they will, I suspect, be of great value both to professional biologists and to the book's intended audience: students and amateur naturalists. I know of no equivalent compilation with coverage ranging through

bacteria to fungi, plants and animals, and both the choice of taxa and the comments in each entry seem appropriate.

But the list of disciplines included, with biochemistry and cell structure and function among them, worries me. How many amateur naturalists want to know about the sliding-filament theory of muscle contraction? And, curiously, though many enzymes have entries to themselves, those that might interest naturalists, such as esterases (for their importance in electrophoretic studies) and ribulose biphosphate carboxylase (the commonest protein in the world) do not appear. Surely natural history is not concerned with mechanisms at the infra-organismic level, and biochemistry and cell biology are out of place in a book such as this. Conversely, the word "history" is used here in the Aristotelian sense, a general account of natural phenomena; so it is right that the earth and atmospheric sciences should be represented.

What, then, of the details? Here I can only turn to the areas I know best. Generally, the definitions seem good, though I would quibble with some (for example those of *fitness*, *polyploidy* and *red light*), and others seem a little dated (under *succession* no mention is made of concepts such as *facilitation*). But in a book of over 12,000 entries it is pointless to pick out one or two individual definitions. More disturbing is the absence of some terms:

mimicry (unless you look up *Müllerian*), CAM (though *C₄ photosynthesis* is in), Red Queen, the MacArthur–Wilson hypothesis (though *island biogeography* makes it), Clements (though *Gleason* gets a mention). There is in this a certain lack of consistency. Some entire fields seem simply to have been overlooked: for example, a whole suite of terms in common use in the study of plant-breeding systems (geitonogamy, gynodioecy, incompatibility) are undefined, at least in this sense. Similarly the coverage of statistical terms is poor and uneven: *correlation* appears, but analysis of variance, chi-squared, degrees of freedom, regression and *t*-test do not.

It is surprising that a dictionary of natural history should seem weakest in its coverage of ecology and evolution, but perhaps this is because ecology, though rich in jargon, has yet to develop a stable vocabulary. I doubt whether this dictionary will enable amateur naturalists to understand the *American Naturalist*, but it contains a wealth of useful taxonomic, geological (in the broad sense) and atmospheric information. Hidden away in it, too, are some real gems: apparently you do get blue moons, particularly in China! □

Alastair Fitter is a Lecturer in the Department of Biology, University of York, York YO1 5DD, UK.

Policing the paper chase

Stevan Harnad

A Difficult Balance: Editorial Peer Review in Medicine. By Stephen Lock. Nuffield Provincial Hospitals Trust, London: 1985. Pp. 172.*

How grave a problem is peer review? The answer is likely to depend on which of the following propositions one judges to be closer to the truth: (i) most of the scientific research that appears in print is significant and essential to the progress of science; (ii) most of it is neither significant nor essential. There is evidence in favour of both propositions. Most papers have minimal content and are rarely or never cited or built upon. On the other hand, it is not clear that high-quality work could be done in a less permissive atmosphere; so, the wheat may always be proportional to the chaff.

Propositions (i) and (ii) have very different implications for the problem of peer review. If most published work is significant, then it is critical what one accepts and rejects for publication, and how: the consequences of the two types of errors, *false negatives* (failures to accept important papers) and *false positives* (failures to reject wrong or trivial papers), loom large. If, instead, most published work is insignificant (which is equivalent to saying that most work is published), then the repercussions of the two types of errors diminish considerably.

What is peer review? In his well-documented and informative survey of editorial procedures in the biomedical sciences, Stephen Lock, editor of the *British Medical Journal*, indicates that it is the relatively recent practice of having work formally (and usually anonymously) evaluated by fellow-scientists. Publication and funding decisions are then made on the basis of the recommendations of these "referees". But whereas the institutionalization of the practice may be recent, it seems apparent that human judgement (not necessarily on the part of peers) has always entered into publication and funding decisions, even in the good old days when just about all research was published and funded.

The funding situation has unfortunately become much more competitive (and perhaps here a more serious problem is developing), but it is still true that just about everything is being published. The

only question is where. Peer review seems to serve to channel the higher quality work towards the more prestigious and widely read journals: in effect, it serves as a selective filter for the beleaguered reader who must keep up with the glut of literature.

Why is there a literature glut? Lock indicates that the emphasis placed on publication-counts in decisions about appointments, promotion and even funding has favoured the strategy of publishing as much as possible, in minimal increments. He sees it as an urgent problem — though not strictly one of peer review — to stem this swelling tide of superfluous publication, perhaps by permitting individuals to publish only five articles per year. This suggestion is not as capricious as it may sound, and would certainly be one way of improving the ratio of quality to quantity in the literature. An intermediate measure would be to place more emphasis on citation-counts than on publication-counts. Unfortunately, however, both of these are merely black-box, quantitative criteria.

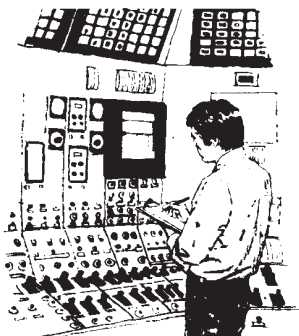
Peer review looks into the black box, but how discerning is it? False positives are easier to identify than false negatives, and Lock surveys some of the prominent cases of error and fraud that have found their way into print. Yet their proportion seems low, probably because science is cumulative and self-corrective, and one cannot build much on a defective foundation (although Lock does give an alarming example of an over-hasty clinical application of a false toxicological report, which had serious consequences for public health). Proposition (ii) — that most published work is inconsequential — suggests another reason why false positives may not represent such a critical problem.

But even proposition (ii) implies that false negatives may be a problem. A conservative tendency (and perhaps a regression towards the average) in collective endeavours puts disproportionate innovativeness at something of a disadvantage among peers, and this is most likely to occur at the high end of the spectrum of significance. Yet quackery and dilettantism need to be filtered out too, and most articles do benefit from the revision they have undergone through peer scrutiny.

False positives and negatives are optimization problems, and Lock offers many useful ideas for investigating and improving the peer-review system. One wonders, however, about the generality and representativeness of conclusions based largely on the biomedical and social sciences. The picture may be quite different in the physical sciences and mathematics (which have lower rejection rates and perhaps different evaluative criteria), and may even vary with whether research is experimental or theoretical, pure or

Proceedings

of Papers presented at
American Nuclear Society
Technical Meetings



write for free
Publications and
Services Catalog.



American Nuclear Society
555 N. Kensington Ave.
La Grange Park, IL 60525
(312) 352-6611

Reader Service No.2

* Available in Europe from STM Distribution Ltd, Enterprise House, Ashford Road, Ashford, Middlesex TW15 1XB, UK, and in North America from ISI Press, Philadelphia. Price is £9, \$24.95.

applied, specialized or interdisciplinary.

Nevertheless, Lock's survey and recommendations are welcome. They include a suggestion that in special cases it would be useful to complement closed peer review with the publication of open peer commentary. I cannot but second this, particularly as 39 of the 281 items in Lock's reference list come from a controversial peer-review study that appeared in the peer-commentary journal I edit — a study whose themes, Lock writes, "run through this monograph as a *leit-motiv*, almost as obsessively (to mix my metaphors) as the Suez canal ran through Lady

Eden's drawing room during the 1956 Suez crisis".

There is an interesting parallel between beliefs about "the peer review system" and beliefs about "the scientific method". Those who favour proposition (i) tend to worry about biases and conspiracies that could impede the ideal march of progress. Those, like myself, who lean towards (ii), feel that it's more a matter of trying to reduce inefficiencies in yet another very human domain of endeavour. □

Stevan Harnad is Editor of The Behavioral and Brain Sciences, 20 Nassau Street, Princeton, New Jersey 08542, USA.

Hormonal activity

Michael J. Brownstein

Neurobiology of Vasopressin. Edited by D. Ganten and D. Pfaff. Springer-Verlag: 1985. Pp. 203. DM 98.

Vasopressin. Edited by Robert W. Schrier. Raven: 1985. Pp. 577. \$104.50.

THE history of research on vasopressin goes back almost a hundred years, to 1895, when Oliver and Schafer showed that mammalian pituitary extracts had pressor (blood-pressure-elevating activity). Three years later Howell demonstrated that the factor responsible resided in the posterior or neural lobe of the gland. Subsequently, Dale found that neural lobe extracts possessed oxytocic activity (that is, they contracted uterine smooth muscle) and von den Velden showed that these extracts could inhibit diuresis as well. In the course of the next 20 years it was established that the oxytocic and pressor-antidiuretic effects were mediated by separate factors, oxytocin and vasopressin, respectively. These were finally purified, characterized and chemically synthesized by du Vigneaud and his colleagues in the 1950s.

Early workers found it difficult to believe that the posterior pituitary could elaborate hormones. Cytologically it is

very different from most glands. The revolutionary discovery by Bargmann and Scharrer in 1951 that the neural lobe is part of a neurosecretory system resolved this problem. The vasopressin in the neural lobe is actually in axons and nerve terminals provided by hypothalamic neurones; vasopressin is released from these nerve terminals into blood vessels to act on its peripheral targets.

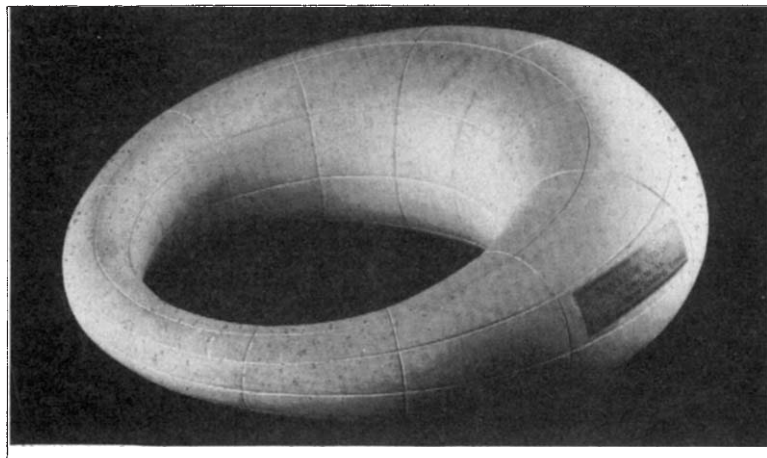
Neurobiology of Vasopressin, the fourth book in a series on current topics in neuroendocrinology, contains five chapters contributed by a total of nine authors. It nicely outlines the research work on vasopressin of the past decade, a period during which many other neuropeptides have come under careful scrutiny. The amino-acid sequence of the vasopressin precursor is now known and its gene has been characterized. Cells that make it have been visualized immunocytochemically. (Many of these lie outside the nuclei that project to the posterior pituitary and seem to be involved in central regulation of autonomic function; surprisingly, some vasopressin-positive cells are present in peripheral tissues such as the adrenal and ovaries.) Steps are being taken towards understanding how vasopressin's synthesis is regulated, and how the afferent inputs to the cells that manufacture it control its production and release in response to changes in water balance.

In spite of the fact that vasopressin's name derives from its effects on blood vessels, these effects are only seen at high dose levels and its cardiovascular effects attracted little interest for many years. This is no longer the case. It is now understood that vasopressin's vasoconstrictor effects are masked by some of its other actions: facilitation of the inhibitory actions of baroreceptive pathways on sympathetic vasomotor activity, and depression of renin secretion. Appropriately, therefore, a large part of *Vasopressin*, the second book reviewed here, is given over to descriptions of the hormone's interactions with central and peripheral baroreceptor systems and regional vascular beds. Also included are sections devoted to the cellular action of vasopressin, vasopressin antagonists, osmoregulation of drinking and vasopressin secretion, neural control of vasopressin release, and the role of vasopressin in producing pathological states.

While many of the individual chapters are very good, the overall organization of the book leaves something to be desired. Those new to the field would have profited from introductory descriptions of vasopressin's structure, biosynthesis, site of production and release, actions and receptors. It is only in the context of this background information that subsequent chapters, which deal with intricate physiological and pharmacological experiments, can be appreciated.

For this reason, I would suggest that readers should tackle *Neurobiology of Vasopressin* before attempting to digest its larger sister volume; in particular, the chapters on volume and cardiovascular regulation in Ganten and Pfaff's book will provide helpful overviews for those attempting to get their bearings on the vasopressin literature. Both books are worth studying, and that neither of them is quite up to date might be seen as a plus point; work on vasopressin is moving along very quickly. □

Michael J. Brownstein is in the Laboratory of Cell Biology, National Institute of Mental Health, Bethesda, Maryland 20205, USA.



ONE of the pleasanter branches of nineteenth-century mathematics was the study of curves and surfaces in space. In order to examine and communicate their properties, geometers built three-dimensional models which frequently possess that interplay of simplicity and complexity, of order and chaos, which characterizes the best abstract art.

The illustration shows one of the "cyclides of Dupin", a class of surfaces whose lines of curvature are all circles or straight lines. Photographs of 132 such models, with mathematical commentary, have been collected into a two-volume work, *Mathematical Models*, edited by Gerd Fischer and published by Vieweg, Wiesbaden, FRG, as part of their two-hundredth anniversary celebrations. In Britain, the books are distributed by Wiley; in the United States, by International Publishing Services, PO Box 230, Accord, Massachusetts 02018.

David Singmaster

FLAMMARION
MÉDECINE - SCIENCES

Jean-françois BACH
immunologie

Immunologie

J.F. BACH et Ph. LESAVRE

De façon claire, directe et très illustrée, les principales connaissances modernes de l'immunologie fondamentale et de l'immunopathologie.

1981, broché, 328 p., 198 fig.,
185.00 FF.

**Jean-François
Jocaut**

**philippe
lejeune**

Immunologie - J.F.

3^e édition de ce livre de

Immunologie - J.F. BACH

3^e édition de ce livre de référence totalement réécrit.
La source la plus actuelle et la plus complète des connaissances en immunologie.

1986, relié, 1100 p., 370 illustr.
697.00 FF.

BIOLOGIE MOLÉCULAIRE DE LA CELLULE

Biologie Moléculaire de la Cellule.

B. ALBERTS, D. BRAY, J. LEWIS,
M. RAFF, K. ROBERTS,
J. WATSON

Traduction française du célèbre ouvrage américain. Rend compte, de façon exhaustive, de toutes les connaissances actuelles sur la Cellule à un niveau très approfondi.

1986, 1300 illustr. en 2 coul.,
édition brochée: 620,00 FF,
édition reliée: 695.00 FF.

**4, rue Casimir-Delavigne
75006 PARIS**

Men and machines remembered

J.M.M. Pinkerton

Memoirs of a Computer Pioneer. By Maurice Wilkes. *MIT Press:1985. Pp.240. \$19.95. £19.95.*

ELECTRONIC computing and communications technologies are of indisputable social and economic importance. To anyone in the field, and many outside it, the history of their development recorded by one of the innovators must be of the greatest interest. Professor Wilkes's book, part of a historical series published by MIT Press, is therefore likely to find a large and receptive audience.

About half of the book deals with the author's family background, his education, his radio research at the Cavendish Laboratory before the war under J.A. Ratcliffe and at the Mathematical Laboratory in Cambridge, and his varied wartime activities as a government radar scientist. All this experience was in fact invaluable, not only because many of the electronic techniques used for radar were readily adapted to the design of early computers, but because Wilkes had become familiar with problems and methods of computation for which computers were suitable.

A fascinating chapter describes the conception, planning and successful construction of EDSAC 1 in the Mathematical Laboratory in the years 1946 to 1949. Wilkes records this in such straightforward terms as almost to belittle the achievement. He carefully acknowledges where due his indebtedness to others for inspiration — to John von Neumann; to speakers on the course at the Moore School of the University of Pennsylvania in 1946, when a variety of techniques were proposed to build a working computer; to the Eckert-Mauchly report proposing a design for EDVAC; and, especially, to Douglas Hartree.

As the series editors say in their foreword, Wilkes aimed not merely to build a machine that would work, but also to investigate programming as an art and an intellectual process applicable to problems, for example in physics, whose numerical solution called for extended computation. While scientists at Cambridge were using the differential analysers, originally developed by Hartree at Manchester and already installed in Wilkes's laboratory, it was anticipated that EDSAC would be faster and better in all respects. With Wheeler and Gill, Wilkes was the first to write subroutines and incorporate them in programs, and a library of subroutines was quickly assembled for programmers' use. Even before EDSAC was

finished, much thought had been given to feeding decimal numbers in and out; for scientific work, the simple paper tape and teleprinter combination adopted was reasonably well matched to EDSAC's arithmetic speed.

Remarkably soon after EDSAC successfully ran programs, Wilkes started to teach others how to develop them, very much in the tradition of the Cavendish Laboratory. Ratcliffe, whose guidance in his earlier career Wilkes acknowledges, used to say that no research was complete until it had been taught, not just published. The course first given in 1950, was the basis of the first book on programming ever to appear — *The Preparation of Programs for an Electronic Digital Computer, with Special Reference to the EDSAC and the Use of a Library of Subroutines* (Addison-Wesley, 1951), which was re-printed by Tomash and MIT Press in 1982.

After starting an informal computing service, based on EDSAC, for the whole University, Wilkes began to design EDSAC 2. This was to be a parallel machine retaining the ultrasonic memory, though in the event core memories were invented in time. A major innovation was microprogramming, a concept Wilkes first described in 1951 and which, of course, has since had a profound influence on computer design.

In the book Wilkes refers to the key contribution he made in the early 1950s to the LEO project (Lyons Electronic Office, the first computer built specifically for clerical work) and later, obliquely, to his present job with Digital Equipment Corporation. But he says little about the guidance he gave as Director of the Cambridge Computer Laboratory until he moved to the United States in 1980, not only to his students but to anyone else who asked for it, as many did.

The book concludes with chapters on more recent developments on both sides of the Atlantic, with many of which Wilkes was involved (for example use of capabilities and the Cambridge Ring LAN). In sum, this is a straightforward personal account, clearly and readably told; it represents a notable contribution to the early history of ideas and practice in digital computing, as they have evolved throughout the world. $\square$

J.M.M. Pinkerton, formerly Manager, Strategic Requirements, with ICL, now runs an information-technology consultancy, McLean Pinkerton Associates, 2 Sandown Road, Esher, Surrey K10 9TU, UK.

● Recently published by Morgan Kaufmann is *Machine Learning: An Artificial Intelligence Approach, Vol. II*, edited by Ryszard S. Michalski, Jaime G. Carbonell and Tom M. Mitchell. Distributor is W.H. Freeman, price is \$39.95, £39.95. The previous volume was reviewed by Margaret Boden in *Nature* **308**, 89 (1984).

High-resolution sedimentary record of a geomagnetic reversal

J.-P. Valet, C. Laj & P. Tucholka

Centre des Faibles Radioactivités, Laboratoire mixte CNRS-CEA, Avenue de la Terrasse, Parc du CNRS, 91190 Gif sur Yvette, France

A detailed study of a geomagnetic reversal shows that during the transition the geomagnetic field undergoes both short- and long-term fluctuations. The first are thought to be linked to the reversal process; the second are present with the same time constants both before and during the reversal. This suggests that the secular variation, at least in its longer-period terms, is not affected by the reversal of the main field. This feature has important consequences for the validity of the frozen-flux approximation.

SEVERAL recent publications have demonstrated that palaeomagnetic studies of reversals may contribute to a better understanding of the mechanisms of the geodynamo. For example, a comparison of the trajectories of the virtual geomagnetic pole (VGP) during the Brunhes-Matuyama transition, as seen from two different sites¹, has led to the conclusion that the geomagnetic field shows a predominantly non-dipole behaviour during a reversal. Other studies have demonstrated that, although axisymmetrical terms certainly play a key role in the morphology of the transitional field, large non-zonal terms are also present²⁻⁴. Various models for reversals have been proposed to account for these observed features^{5,6}.

With a few exceptions⁷⁻⁹, however, the published records are not sufficiently accurate to provide detailed information about the precise geomagnetic variations occurring during a reversal, often because of a lack of intermediate directions, or because of bad stratigraphic control of the different samples. Here we describe the results obtained from a reversal recorded in a sequence of Tortonian marine clays near the village of Skouloudhiana in the Khania province of western Crete. We have already described^{10,11} some of the main characteristics of the different reversals recorded in this section and in another nearby section at Potamida. Here we consider only the upper reversal in Skouloudhiana (KS 06 in the previous work), which has been studied in great detail. About 100 additional transitional palaeomagnetic directions have been obtained and the results analysed with suitable mathematical techniques.

Geological setting and sampling

The section at Skouloudhiana comprises 54 m of homogeneous blue-grey marine clays. In the sampled portion of the section, which is well-exposed, there is no evidence of faults or slumped beds. The bedding plane is very nearly horizontal, so no tilt correction has been applied to the results. Previous biostratigraphic results have established that the section belongs to the Upper Tortonian stage¹², and correlation with the geomagnetic polarity timescale¹² gives an age of 5.87 Myr for the upper R-N (reverse- to normal-polarity) sixth reversal (upper part of epoch 6).

In a dense sampling programme during two field trips, 295 cores (6-15 cm long) were collected over a stratigraphic height of 4.7 m, comprising the entire transition zone and also levels outside the zone. The cores were cut in the field into 22-mm samples.

Very careful attention was paid to the determination of the exact stratigraphic position of the samples. Stratigraphic reference levels were marked every ~20 cm on the surface of the section, and the position of each core was measured with respect to these levels. Measurements of core positions with respect to different reference levels yielded identical results to within 0.5 cm: this value was then taken as an estimate of the error. Precise stratigraphic marks and photographic records have

allowed us to correlate the stratigraphic position of cores drilled during the second field trip with those obtained during the first, with almost the same accuracy. The final stratigraphic position of the different samples obtained from the cores was calculated using a computer program which took into account the usual orientation parameters of the cores, their stratigraphic position, and even the width of the diamond saw used to cut the samples.

The result of the sampling programme was an almost continuous stratigraphic succession with several specimens at each horizon, so that different experiments could be performed on samples from the same stratigraphic level.

Sedimentology and magnetic mineralogy

Although the good magnetic properties of the clays of the Skouloudhiana section had already been established¹², further analyses were made in the stratigraphic interval corresponding

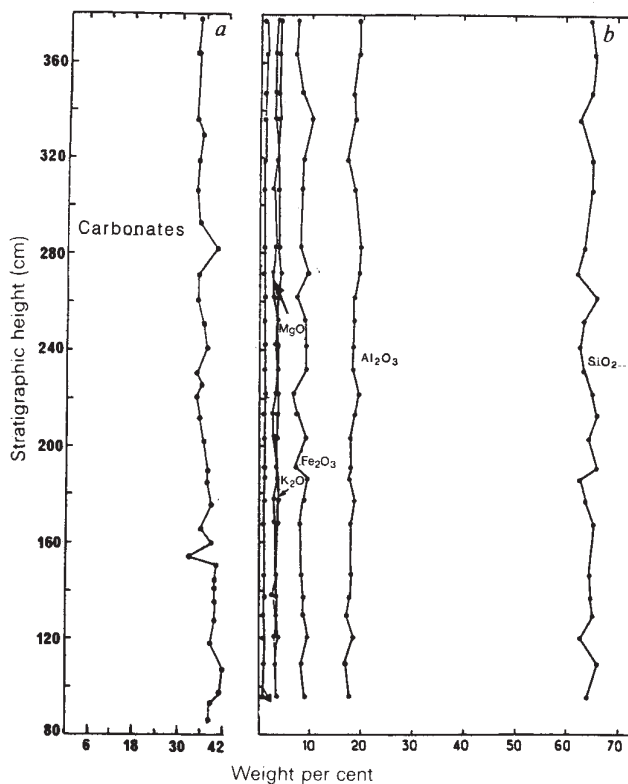


Fig. 1 Evolution of characteristic sedimentary parameters as a function of stratigraphic height. *a*, Variations of the carbonate fraction expressed in percent of the total weight; *b*, variations of the major elements expressed in weight per cent of the residual carbonate-free fraction.

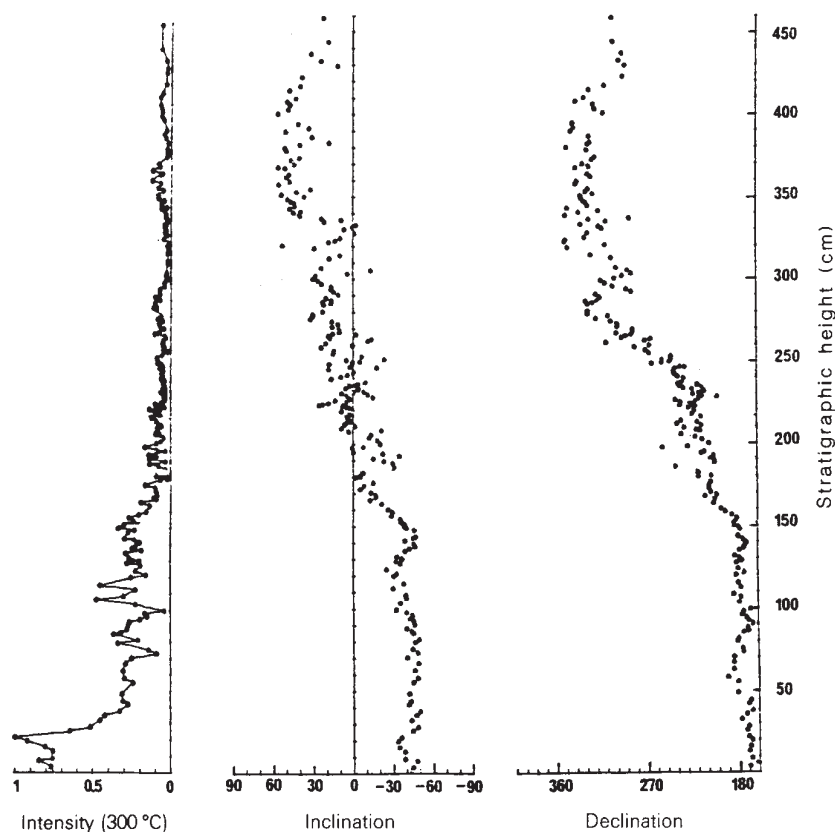


Fig. 2 Records of the intensity (after heating at 300 °C), declination and inclination, plotted against stratigraphic height. The results have been averaged for 10 pairs of samples separated by less than 0.5 cm.

to KS 06. Particular attention was paid to those parameters whose possible variation with stratigraphic height could be an indication of large perturbations or changes in the sedimentary regime.

At 34 levels regularly spaced over the KS 06 interval, X-ray diffraction analyses show the mineralogy to be dominated by detrital quartz, clay minerals and some feldspar, in addition to the carbonate fraction. Illite, kaolinite and chlorite have been identified as the main components of the clay fraction. Both the nature and the relative abundances of the different minerals are quite constant throughout the sampled interval. The carbonate content (Fig. 1a) shows no significant deviation from a mean value of 37 ± 2.5 wt%. Furthermore, the relative abundances of major elements (Fig. 1b) in the decarbonated fraction, determined by microprobe analysis, are also remarkably constant over the KS 06 transition zone.

The nature of the magnetic minerals was investigated by combining measurements of the IRM (isothermal remanent magnetization) acquired step-wise in fields up to 1.5 tesla (T) and of the low-field magnetic susceptibility χ . The results obtained on a series of 40 samples regularly spaced over the sampled zone indicate that $90 \pm 4\%$ of the saturation IRM (SIRM) is acquired in a field of 0.2 T. Successive alternating-field or thermal demagnetizations show that the median destructive field of the SIRM is 25 mT and that $>95\%$ of it is lost by heating at 580 °C, complete removal being obtained between 580 and 630 °C. There are no significant differences in the low-field susceptibility of the samples, indicating that the concentration of magnetic minerals does not vary appreciably over the transition zone. Moreover, the χ /SIRM ratios are grouped around the value $4 \times 10^{-5} \text{ m A}^{-1}$, typical of fine-grained magnetite^{13,14}. All of these experiments indicate that the remanent magnetization of the clays in the KS 06 transition zone is mainly carried by magnetite, with a mineral of higher coercivity, presumably haematite, also present in very small amounts.

These magnetic minerals proved to be stable during heating. Indeed, the low-field magnetic susceptibility measured at each

step of the thermal demagnetization shows only moderate changes (up to a factor of 3) in the interval between 450 and 550 °C.

Measurements of the anisotropy of magnetic susceptibility in a series of samples distributed over the transition zone show that the magnetic fabric of the sediment is characterized by an oblate ellipsoid with a very weak magnetic lineation (1.015). The axes of minimum susceptibility $K_{\min}$ are very close to the vertical while those of maximum susceptibility $K_{\max}$ are horizontal and their directions are largely scattered in the foliation plane. These results are a good indication of sedimentation in rather calm water; distortion of the results by bottom currents, if any, is certainly very limited.

All of the experiments carried out in the mineralogy and sedimentology of the samples section provide evidence of a uniform, undisturbed sedimentary environment. We are well aware that this evidence does not constitute absolute proof of a constant sedimentation rate during the entire interval of sediment deposition. However, the fact that none of the measured parameters shows significant time variation makes this hypothesis very reasonable, and we can rule out the possibility of any major changes in the sedimentation rate.

Transitional palaeomagnetic directions

Measurements were made partly with a Digico spinner magnetometer and partly with a LETI cryogenic magnetometer. Both alternating-field and thermal demagnetization were used on pilot samples. The results obtained by thermal demagnetization were systematically more consistent than those obtained with the alternating field; for this reason, thermal demagnetization was used throughout, with steps of 40–50 °C from 100 °C to the highest temperature at which the measurements could still be made with sufficient accuracy.

The final palaeomagnetic directions were determined from demagnetization plots by hand-fitting a straight line through at least the last 4–5 points. Two secondary components have been identified. The first one is completely removed at ~ 100 –150 °C.

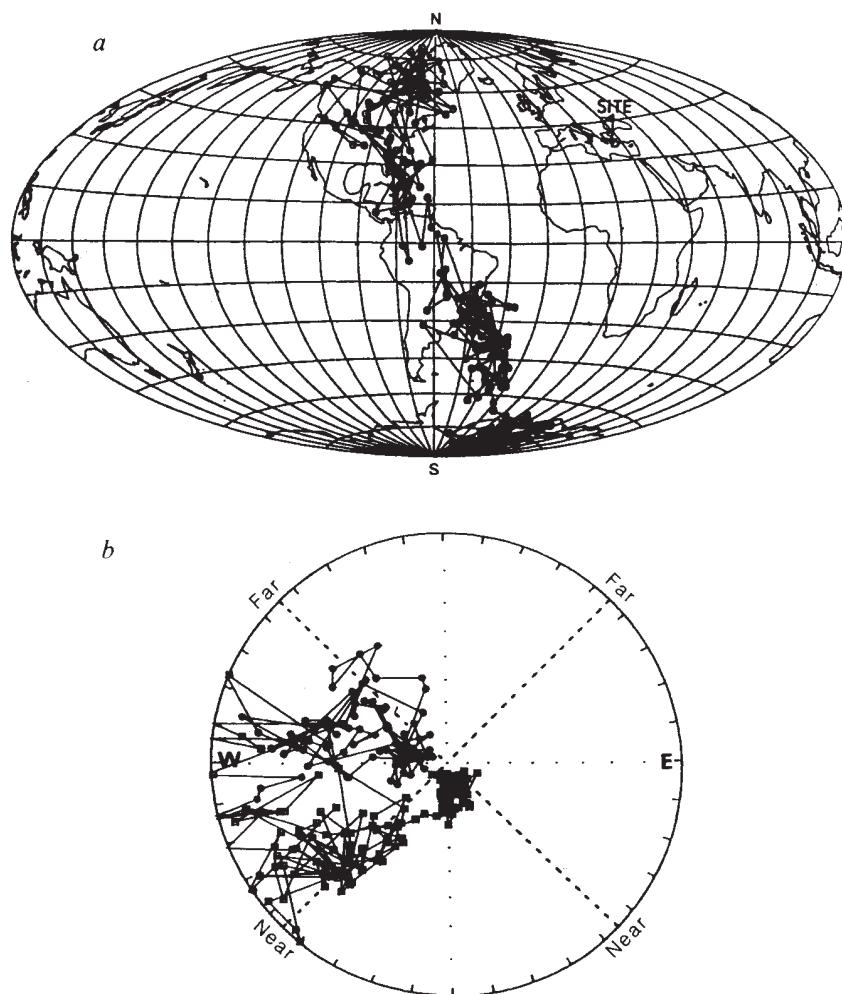


Fig. 3 *a*, VGP path relative to the KS 06 transition calculated from the results in Fig. 2 and plotted in the planisphere centred at 300°E (83° west from the site longitude). *b*, Polar stereographic projection of the 'rotated directions'²⁰ D' , I' .

The second one, clearly observed in reverse- and intermediate-polarity samples, with a direction similar to that of the present geomagnetic field, is probably of viscous origin and is removed at 200–300 °C. No other unstable components have been found up to the highest temperatures, so that the final palaeomagnetic directions could be determined with a good overall accuracy (only 11 out of 245 demagnetized samples did not yield a stable direction and were rejected).

Figure 2 shows the results obtained after demagnetization, in the form of a detailed record of inclination (I) and declination (D) as a function of stratigraphic height. Each point corresponds to the result obtained on a single sample, except for 10 pairs of samples whose stratigraphic height was found to differ by less than 0.5 cm. Because this difference is less than the estimated accuracy of the measurement of the stratigraphic position, the results from these 10 pairs were averaged, and only one point is reported for each of them; the mean dispersion of the averaged directions is 11°.

The inclination values measured on samples from outside the transition zone are shallower by ~10° than the value of 54.5° expected at the site latitude, assuming an axial geocentric dipole field. This inclination anomaly has been frequently observed in similar sediments from throughout the Hellenic Arc¹⁵. Although no definite conclusions have been reached concerning its origin, an overall analysis of the results from many sections suggests that compaction of the sediments is a plausible cause.

The duration of the transition can be estimated by considering the stratigraphic height relative to the transitional directions and assuming a regular accumulation rate. The boundaries of the

transition, defined as where the directional changes exceed the amplitude of the field variations before and after the transition, show that the stratigraphic height occupied by intermediate directions is 1.80 m. This value is twice as large as for records of the same transition in the adjacent sections Potamida 1 and 4, although the lengths of the polarity zones deduced from magnetostratigraphic results^{12,16} are the same in the three sections. We therefore conclude that the sedimentation rate during this reversal was higher than the mean value for the whole section and, moreover, that the deposition was roughly twice as fast in Skouloudhiana as in Potamida at that time. The average sedimentation rate determined by Langereis^{12,16} for western Crete is ~4 cm per 1,000 yr. Using a value twice as large for the KS 06 reversal in Skouloudhiana we obtain a duration of the reversal of ~20,000 yr. Estimates of the transition duration from directional changes^{18,19} range from 2,000 to 10,000 yr.

Discussion

Figure 3 presents the results in the alternative forms of the VGP path (Fig. 3*a*) and the 'rotated directions'²⁰ D' , I' (Fig. 3*b*); the two representations yield very similar results. The record is characterized by a considerable number of intermediate directions—more than 120 VGP positions between 60°S and 60°N. A very detailed record has thus been obtained and it is possible to observe some characteristics of the reversal. Some of them, previously mentioned elsewhere^{11,12}, will not be discussed here, and we will concentrate mainly on the detailed variations of the transition field.

Before proceeding, it is important to examine whether the

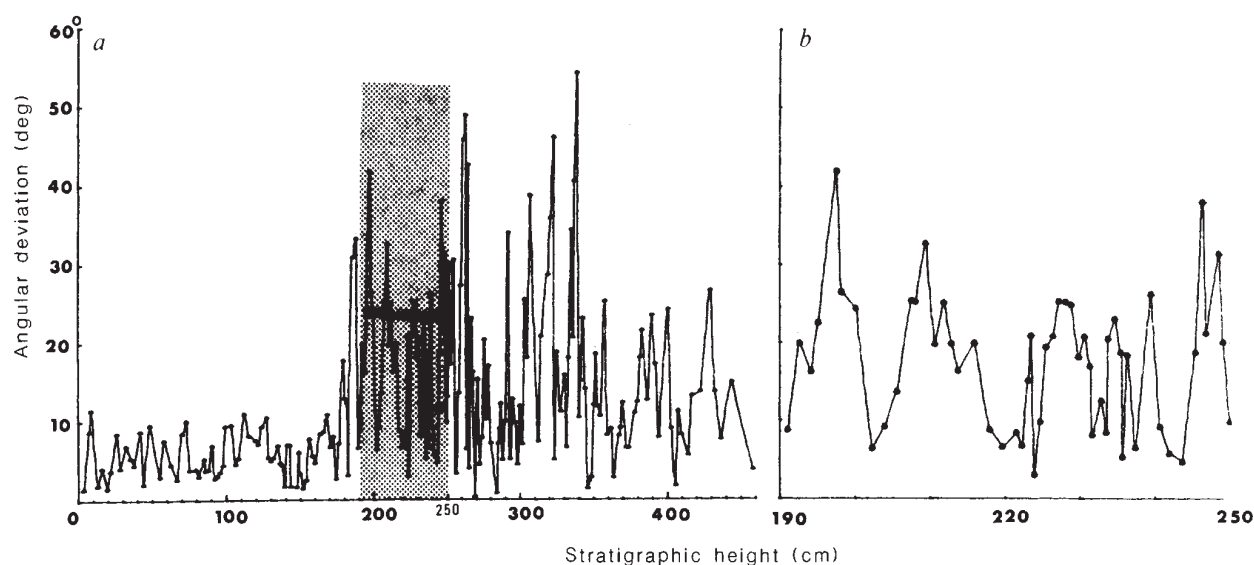


Fig. 4 *a*, Angular deviations between the successive magnetization vectors as a function of stratigraphic height. The shaded part of the plot is enlarged in *b*, to show that some of the fluctuations include several successive points.

sediment of KS 06 is adapted to this kind of study. It is well known that in a newly deposited sediment, some grains can remain free to rotate inside water-filled voids after deposition, so that a range of blocking times will be associated with each horizon within the sample, thus introducing additional spread into the magnetic record (a phenomenon known as post-detrital remanent magnetization, PDRM). Experiments using laboratory-deposited sediments^{20,21} have shown that this process depends on the type of sediment. Following preliminary work by Salloway²², the results of a model simulating the remanence acquisition in a sediment column deposited during a reversal show that a very long acquisition delay eliminates any rapid field fluctuation. Consequently, any fluctuation superimposed on a VGP trajectory is representative of a real geomagnetic field variation and a reasonable estimate of the depth of smoothing introduced by the PDRM process can be deduced from the length of the oscillations.

The results obtained on KS 06 show the presence of oscillations in the geomagnetic record, which in addition to their physical significance prove that the smoothing due to PDRM is very limited. Two types of oscillations are observed, with different time constants.

Rapid fluctuations. Fluctuations occur on a very short timescale, and a plot of the angular deviation between two successive vectors as a function of stratigraphic height (Fig. 4*a*) shows a significant increase of the fluctuation amplitude during the transition. Similar fluctuations are frequent in most sedimentary records of reversals and have usually been implicitly assumed to represent an overall random noise, although no thorough investigation of their origin has been performed. In the case of KS 06 several features indicate that no factor, other than the geomagnetic field itself, can account for the amplitude of the observed fluctuations.

First of all, note that several fluctuations are characterized by several successive points (Fig. 4*b*), and thus cannot be entirely related to random errors. The instrumental error in the measurements does not exceed 5° for the least magnetized samples, and the demagnetization diagrams are of a generally good quality. Thus, variations as large as 40–60° cannot be explained by inaccuracies in the measurements or in the determination of the final palaeomagnetic directions. Fluctuations can also be induced by errors in the stratigraphic positions of the samples. However, this error is too small (0.5 cm) to account for the amplitude of the observed fluctuations. Finally, Merrill and

McElhinny²³ have pointed out that “a secondary component acquired in the relatively high fields in non-transitional times can produce a large spurious direction when vectorially added to a primary direction acquired in a relatively weak transition field”. Poor magnetic cleaning could thus result in fluctuations not related to the reversal process. If this were the case, however, the vector differences between the palaeomagnetic directions of successive samples should be predominantly aligned with the direction of the dipole field at the site (or its opposite for secondary components acquired in a reverse field). A stereographic plot of these vector differences, shown in Fig. 5, shows no evidence of such a grouping, indicating that secondary components certainly play no major role in the observed fluctuations.

Thus we think that at least a significant part of these fluctuations represent very rapid variations of the geomagnetic field, and therefore there has been negligible smoothing due to PDRM. Of course, we do not rule out the effects of the previously mentioned parameters, it would not be reasonable to attribute a quantitative value to the amplitude of the fluctuations; neither do we propose a time constant, because of the averaging introduced by the stratigraphic height covered by each sample.

In their interpretation of the volcanic record of the Steens Mountain reversal, Prevot and co-workers^{3,8} suggested that the large directional jumps could reflect an increase in the level of turbulence within the Earth's liquid core during reversals. The significant increase in the amplitude of the variations observed in the KS 06 record could also reflect such an increase in turbulence.

Long-term variations. Fluctuations also occur on a considerably longer timescale, but are masked by the high-frequency fluctuations, so that a smoothing of the data was necessary to reveal them. Three independent smoothing methods have been used, involving running averages with gaussian windows, and polynomial and cubic spline approximations. The results in Fig. 6*A*, obtained after smoothing by cubic splines with 27 knots, are virtually identical to the results of the other smoothing procedures. The directional variations display a succession of regular oscillations, which provide additional proof that the PDRM is extremely limited.

These slow fluctuations are thus truly representative of geomagnetic changes, and therefore merit further consideration. It can be seen in the declination and inclination (Fig. 2), and in the angular variations in Fig. 6*A* (curve *a*), that three well-defined oscillations occurred before the transition, with identical

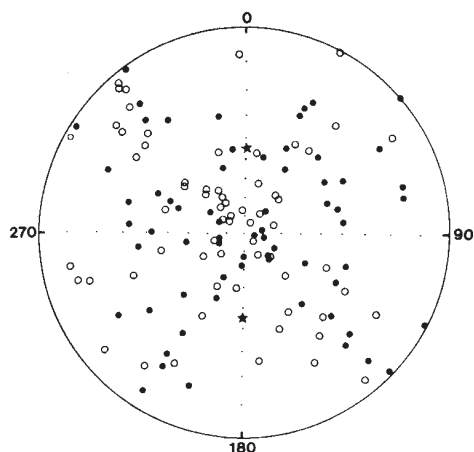


Fig. 5 Stereographic projection of the vector differences between the successive transitional directions. The normal and reverse dipole field directions at the site are represented by asterisks. Symbols, positive inclinations; open symbols, negative inclinations.

wavelengths suggesting a periodicity between 2,000 and 4,500 yr by comparison with the duration of the transition. These values are in the same range as the periodicities which are commonly reported for the non-dipolar field variations obtained from high-deposition-rate lake records during the past 15,000 yr (ref. 24). Variations with similar time constants persist during the transition, and a more quantitative picture is shown by the curve in Fig. 6B, obtained by subtracting from Fig. 6Aa the monotonic angular changes of the reversal (Fig. 6Ab). The results show a very regular succession of oscillations, and the increasing amplitude of these fluctuations during the transition is consistent with the fact that the dipole field intensity decreases. Some characteristics of the non-dipolar geomagnetic field are thus not affected by the polarity change.

Reversal dynamics. The samples are almost equally spaced in stratigraphic height, and the previously described experiments have shown that there is no reason to suspect major changes in the sedimentation rate. In spite of this, it can be seen that the variations (curve *b*) obtained in Fig. 6A after subtracting the oscillating part (Fig. 6b) are not strictly monotonic. This feature is also apparent from the VGP path (Fig. 3a), in which the successive positions of the VGPs are not equally distributed along the trajectory. This inhomogeneous distribution reflects the fact that the reversal did not occur at constant speed. After a quick passage of the pole through the high southern latitudes there is a quiet episode in the mid-southern latitudes, after which the equatorial zone is crossed rather quickly. It can therefore be suggested that during this reversal the least stable 'state' of the field corresponds to the middle of the transition (when the field intensity is very weak^{25,26}). Unfortunately, the number of detailed geomagnetic records is too limited (or the time sequence is too often poorly defined or interrupted) to compare the dynamical characteristics of KS 06 with other distinct reversals. It seems, however, that the VGP path configuration is similar to that associated with an R-N reversal recorded from the Chugwater Triassic formation²⁷, but differs from that reported by Liddicoat⁹ for the Gauss-Matuyama reversal from Searles Valley. There is, of course, no reason why all reversals should be characterized by the same dynamical structure.

Finally, it can be seen in Fig. 6a that the end of the record is characterized by a movement of the VGPs towards equatorial latitudes, which probably reflects a geomagnetic excursion occurring after the reversal itself. Excursions occurring before or after a transition have been observed in some records and

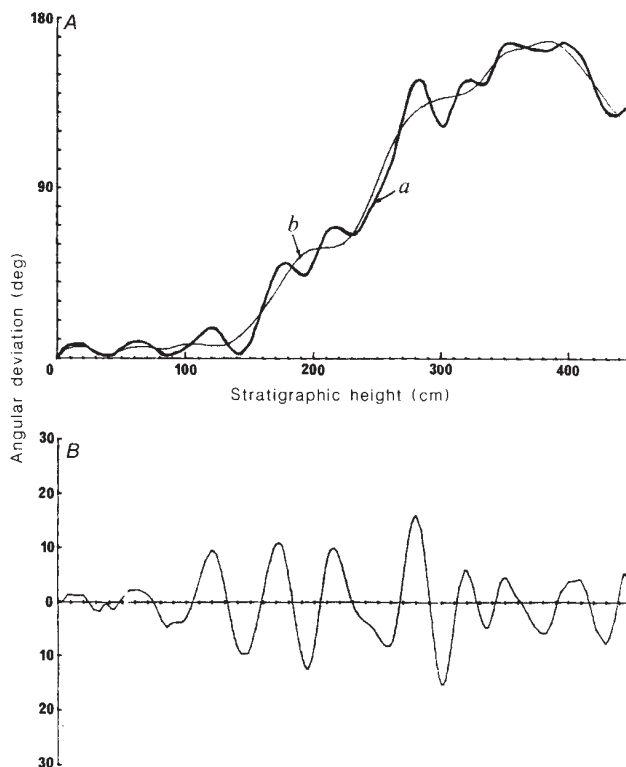


Fig. 6 A, Angular deviations from the pre-reversal direction plotted against stratigraphic height: *a*, after smoothing by cubic splines with 27 knots; *b*, after smoothing with 15 knots. B, Periodic oscillations obtained by subtracting curve *b* from curve *a* in *a*.

they could be related to the decrease or recovery phases of the dipole field. However, it is not always easy to distinguish, as in the case of the Steens Mountain reversal record²⁸⁻³⁰, to what extent these excursions are connected to the reversal. The configuration of transitional and excursive VGPs of KS 06 along the same longitudinal band (see Fig. 2) suggests a similarity of the mechanisms for both and, in this case, this observation can also be reconciled with Hoffman's suggestion³¹ that excursions are aborted reversals.

Conclusion

The high time-resolution of this record, which is less than a few hundred years, has allowed us to observe some detailed characteristics of the transitional field. Apart from the dynamical aspect of the reversal, which could be different for each reversal, the most important result is the presence of two types of geomagnetic field variations during the transition. Very fast fluctuations seem to be directly related to the reversal process and, as has been previously suggested², they could result from an increase of turbulence in the Earth's core during the transition. On another timescale, this record has shown the persistence of long-term variations of the non-dipole field during the transition. The non-dipole field, at least in its longer-term components (of the order of 10^3 yr), has not been affected by the reversal of the main field.

This last feature has some bearing on the validity of the frozen-flux approximation. Le Mouél³² has shown that an outer-core geostrophic flow accounts for the observed temporal variations of the main field (secular variation), that is, for the different time constants of its dipolar and nondipolar parts. Indeed, this motion involves the two parts of the field in a different way and the axial dipolar component is not re-engaged in the process that builds up the secular variation. Assuming furthermore that the frozen flux approximation holds for time

spans of thousands of years, Le Mouél has demonstrated that the non-dipole field cannot disappear during the process of a geomagnetic reversal. On the other hand, Bloxham and Gubbins³³ have found evidence for flux diffusion in the South Atlantic region and have tentatively concluded that the frozen-flux hypothesis can be rejected with 95% confidence. However, as pointed out by the authors themselves, this conclusion depends critically on error estimates for the field at the core-mantle boundary which are difficult to assess, so that at present it cannot be considered as definitely established.

In this respect the observed continuity of the secular variation during the KS06 reversal is to our best knowledge the first experimental evidence supporting Le Mouél's theoretical results.

Moreover, the fact that the processes leading to the observed secular variation would persist unchanged during a reversal suggests that only the internal part of the core would be involved in the mechanism of reversal.

We thank Dr C. Langereis and Dr H. Zijdeveld for many helpful discussions and comments; C. Jehanno for supplying the microprobe analyses; L. De Seze for help in the mathematical treatment by cubic splines; and C. Kissel and S. Sen for help in sampling. The director of the Institute of Geology and Mining Research (IGME) at Athens kindly provided the necessary permits. This work was partly supported by the French CNRS ATP Géologie et Géophysique des Océans and CNRS-INSU ATP-Noyau. This is CFR contribution no. 761.

Received 3 February; accepted 7 May 1986.

1. Hillhouse, J. W. & Cox, A. *Earth planet. Sci. Lett.* **29**, 51-64 (1976).
2. Williams, I. & Fuller, M. *J. geophys. Res.* **87**, 9408-9418 (1982).
3. Prévot, M., Mankinen, E. A., Grommé, C. S. & Coe, R. S. *Nature* **316**, 230-234 (1985).
4. Theyer, F., Herrero-Bervera, E. & Hsu, U. *J. geophys. Res.* **90**, 1963-1982 (1985).
5. Hoffman, K. A. *Earth. planet. Sci. Lett.* **44**, 7-17 (1979).
6. Kaiser, A. D. & Verosub, K. L. *Geophys. Res. Lett.* **12**, 777-780 (1985).
7. Clement, B. M. & Kent, D. V. *J. geophys. Res.* **89**, 1049-1055 (1984).
8. Prévot, M., Mankinen, E. A., Coe, R. S. & Grommé, C. S. *J. geophys. Res.* **90**, 10417-10448 (1985).
9. Liddicoat, J. C. *Phil. Trans. R. Soc. A* **306**, 121-128 (1982).
10. Valet, J.-P., Laj, C. & Langereis, C. G. *Nature* **304**, 330-332 (1983).
11. Valet, J.-P. & Laj, C. *Nature* **311**, 552-555 (1984).
12. Langereis, C. G., Zachariasse, W. J. & Zijdeveld, J. D. A. *Mar. Micropaleont.* **8**, 261-282 (1983-84).
13. Harstra, R. L. *Geophys. J. R. astr. Soc.* **71**, 477-495 (1982).
14. Dankers, P. H. M. thesis, State Univ. Utrecht (1978).
15. Laj, C., Jamet, M., Sorel, D. & Valente, J.-P. *Tectonophysics* **86**, 45-67 (1982).
16. Langereis, C. G. thesis, State Univ. Utrecht (1984).
17. Jacobs, J. A. *Reversals of the Earth's Magnetic Field* (Hilger, Bristol, 1984).
18. Hoffman, K. A. *Rev. Geophys. Space Phys.* **21**, 614-620 (1983).
19. Hoffman, K. A. *J. geophys. Res.* **89**, 6285-6292 (1984).
20. Tucker, P. *Geophys. J. R. astr. Soc.* **63**, 149-163 (1980).
21. Tucker, P. in *Geomagnetism of Baked Clays and Recent Sediments* (eds Creer, K. M., Tucholka, P. & Barton, C. E.) **3** (Elsevier, Amsterdam, 1983).
22. Salloway, J. C. thesis, Univ. Edinburgh (1983).
23. Merrill, R. T. & McElhinny, M. W. (eds) *The Earth's Magnetic Field* (Academic, London, 1983).
24. Creer, K. M. & Tucholka, P. *Phil. Trans. R. Soc. A* **306**, 87-102 (1982).
25. Valet, J.-P. & Laj, C. *Earth planet. Sci. Lett.* **54**, 53-63 (1981).
26. Valet, J.-P. thesis, Univ. Paris XI (1985).
27. Herrero-Bervera, E. & Helsley, C. E. *J. geophys. Res.* **88**, 3506-3522 (1983).
28. Valet, J.-P., Laj, C. & Tucholka, P. *Nature* **316**, 217-218 (1985).
29. Grommé, C. S., Mankinen, E. A., Prévot, M. & Coe, R. S. *Nature* **318**, 487 (1985).
30. Valet, J.-P., Laj, C. & Tucholka, P. *Nature* **318**, 487-488 (1985).
31. Hoffman, K. A. *Nature* **294**, 67-68 (1981).
32. Le Mouél, J.-L. *Nature* **311**, 734-735 (1984).
33. Bloxham, J. & Gubbins, D. *Geophys. J. R. astr. Soc.* **84**, 139-152 (1986).

Cloning the gene for an inherited human disorder—chronic granulomatous disease—on the basis of its chromosomal location

Brigitte Royer-Pokora*, Louis M. Kunkel†, Anthony P. Monaco†, Sabra C. Goff*, Peter E. Newburger‡, Robert L. Baehner*, F. Sessions Cole§, John T. Curnutte|| & Stuart H. Orkin*¶

* Division of Hematology-Oncology, The Children's Hospital, Dana-Farber Cancer Institute, Department of Pediatrics, † Division of Genetics, The Children's Hospital, Department of Pediatrics and the Program in Neuroscience, § Division of Cell Biology, The Children's Hospital, Department of Pediatrics, Harvard Medical School, Boston, Massachusetts 02115, USA

‡ Division of Pediatric Hematology, Department of Pediatrics, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA

|| Division of Pediatric Hematology, Department of Pediatrics, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA

¶ Howard Hughes Medical Institute, Children's Hospital, Boston, Massachusetts 02115, USA

The gene that is abnormal in the X-linked form of the phagocytic disorder chronic granulomatous disease has been cloned without reference to a specific protein by relying on its chromosomal map position. The transcript of the gene is expressed in the phagocytic lineage of haematopoietic cells and is absent or structurally abnormal in four patients with the disorder. The nucleotide sequence of complementary DNA clones predicts a polypeptide of at least 468 amino acids with no homology to proteins described previously.

MANY inherited disorders in man result from mutations in genes whose protein products are as yet unknown. For some disorders specific proteins may be suspected as candidates for the primary gene products, whereas in other situations no biochemical clues exist. Molecular cloning offers an approach to the analysis of disease gene loci which, in principle, may be executed without specific information regarding the nature of the proteins involved. The use of genetic linkage to establish the location of a disease gene within the chromosome complement has generally been considered an initial step in this analysis¹. Major diseases

of unknown aetiology such as Duchenne muscular dystrophy (DMD), Huntington's disease and cystic fibrosis, have been the focus of intensive study from this perspective²⁻⁷. It has been anticipated that extended genomic cloning⁸, coupled with searches for transcribed regions, will provide a means of defining each disease locus in precise genetic terms. Here we describe the successful application of this general strategy to the identification and characterization of the gene involved in chronic granulomatous disease (CGD), an X-chromosome-linked disorder of phagocytic cells. Affected individuals have greatly impaired host defences against infection with a wide variety of microorganisms⁹. In contrast to the normal situation, phagocytes

¶ To whom correspondence should be addressed.

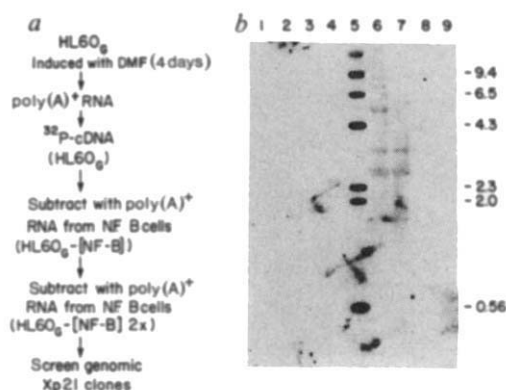


Fig. 1 Detection of a transcribed region of Xp21. *a*, Strategy for the preparation of an enriched cDNA probe. Granulocytic HL60 cells (HL60_G) were used as a source of mRNA and total ³²P-cDNA. Sequences in common between HL60_G and the B-cell line of patient N.F. were removed by the method of Davis³². *b*, Southern blot hybridization³³ of the enriched cDNA probe with *Eco*RI + *Hind*III-digested DNAs of bacteriophage derived from Xp21. Bacteriophage clone designations: lane 1, 469-7A2; 2, 378-40A1; 3, 378-29A1; 4, 378-19A1; 6, 379A6; 7, 379π; 8, 55π; 9, 145-A1. These bacteriophage correspond to clones hybridizing with the pERT clones 469, 378, 379, 55 and 145 (ref. 15). Lane 5, *Hind*III-digested λ DNA as a marker. Hybridization to fragments of 2.5 and 3.3 kb was observed for the samples in lanes 6 and 7. These clones overlapped the same genomic region. Bands of greater size, visible in lane 6, are due to slightly incomplete digestion of phage 379-A6 DNA and not additional hybridizing fragments, as verified by subsequent direct analysis of the phage with cloned cDNA (not shown).

Methods. *a*, Total poly(A)⁺ RNA was prepared from HL60_G cells as described elsewhere²⁰. ³²P-labelled cDNA was synthesized using random primers and AMV (avian myeloblastosis virus) reverse transcriptase³² (specific activity of 1×10^8 c.p.m. μg^{-1}). cDNA was depleted of template RNA by base hydrolysis and hybridized in 0.5 M sodium phosphate buffer, 0.1% SDS, 5 mM EDTA in sealed glass capillaries at 69 °C to a R₀t of 2,700 with a 20-fold excess of poly(A)⁺ RNA prepared from cultured B cells of patient N.F. Single-stranded cDNA was isolated in 0.12 M sodium phosphate on hydroxyapatite at 60 °C and rehybridized to N.F. poly(A)⁺ RNA as described above³². *b*, Single-stranded cDNA (1×10^6 c.p.m.) from the second hydroxyapatite step was hybridized with the digested phage DNAs on a GeneScreen filter in a 1.0-ml solution containing 1 M NaCl, 2× Denhardt's solution, 10% dextran sulphate, 1% SDS, 0.2 M sodium phosphate, 50 $\mu\text{g ml}^{-1}$ salmon sperm DNA, 50 $\mu\text{g ml}^{-1}$ oligo(dA)₁₈ for 60 h at 68 °C. The filter was washed in 1× SSC, 1% SDS, and 0.5× SSC, 1% SDS at the same temperature. Autoradiographic exposure was for 10 days with an intensifying screen.

(granulocytes, monocytes and eosinophils) isolated from the blood of CGD patients fail to generate superoxide and activated oxygen derivatives on ingestion of microbes. The precise lesion in the membrane-associated NADPH-oxidase system of these cells is unknown. Cell complementation¹⁰ and family studies⁹ have defined X-linked and autosomal recessive varieties of the disorder. In most families the disease is transmitted as an X-linked trait. Over the past two decades several proteins, including an unusual *b*-type cytochrome¹¹ and a flavoprotein¹², have been considered as potential candidates for the primary products of CGD loci, but difficulties in solubilizing membrane-associated components have impeded biochemical characterization of the oxidase system. A consistent finding in the X-linked variety, however, is the absence of the haem spectrum derived from cytochrome *b*¹¹. Given the potential complexity of the oxidase system and the possible requirement for assembly of several components into an enzymatically active unit, this recognized deficiency in the cytochrome *b* spectrum may be secondary to the primary genetic abnormality.

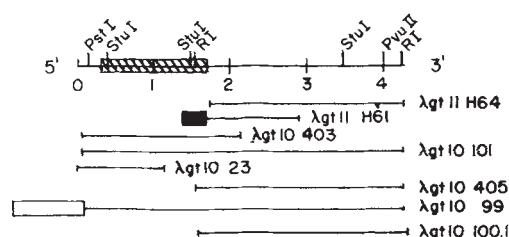


Fig. 2 cDNA clones spanning the 379-encoded RNA. Granulocytic HL60 cDNA libraries were constructed in λgt11 and λgt10 using procedures described elsewhere¹⁸⁻²⁰. Initial screening with a subclone from the 379-A6 bacteriophage yielded λgt11 H64 and H61. The latter contained a segment absent from other overlapping clones (indicated by the solid box) and thought to represent intron sequences. Repeated screenings with additional cDNA subfragments of the λgt10 library yielded cDNAs spanning the entire mRNA, representative examples of which are shown. The open box at the 5' end of clone 99 contained sequences unrelated to the 379 transcript. The cross-hatched box represents the major open reading frame of the mRNA, as defined by DNA sequence analysis (Fig. 6a).

In view of the problems inherent in conventional biochemical analysis, we have adopted a genetic approach to define the molecular basis of the X-linked form of CGD as a means of delineating at least one critical protein of the oxidase system.

General strategy

Four discrete steps are involved in our analysis: (1) The position of the CGD gene on the X chromosome (X-CGD) was mapped by both deletion and formal linkage analysis, described recently¹³. (2) Messenger RNA transcripts derived from the assigned chromosomal region were sought. (3) The relevance of one specific mRNA transcript to the disorder was investigated by studying affected patients. (4) The protein product encoded by the X-CGD gene was predicted from isolated complementary DNA clones. Taken together, the results provide persuasive genetic evidence that we have identified the X-CGD gene and characterized its mRNA transcript.

Expressed Xp21 sequences

From the study of two patients (B.B. and N.F.) with interstitial deletions of Xp21 and from linkage of X-CGD with two Xp21 DNA markers (p745 and pERT 84), the CGD locus has been assigned to Xp21.1 (refs 13-15, 42). This delimits the X-CGD locus to ~0.1% of the genome, or perhaps 3,000 kilobases (kb) of DNA. As part of chromosome walking experiments originally aimed at characterizing the DMD locus, a series of bacteriophage clones derived from seven regions of Xp21 that were absent from the DNA of patient B.B. were obtained^{3,15}, using pERT clones 55, 84, 87, 145, 378, 379 and 469 as probes (L.M.K. and A.P.M., unpublished data). (These bacteriophage clones span at most 10% of the DNA deleted in patient B.B.) In this manner, approximately 250 kb of Xp21 DNA was isolated.

The experiment outlined in Fig. 1a was performed to search for regions of these bacteriophage clones that are transcribed in phagocytic cells. Briefly, a Southern blot filter bearing restriction enzyme-digested bacteriophage DNAs was hybridized with a radioactive cDNA probe enriched for sequences expressed in phagocytic cells. As a source likely to contain transcripts of the CGD gene, we chose cultured human HL60 leukaemic cells¹⁶ treated with dimethylformamide (DMF) for 4-7 days. Such treatment induces the NADPH-oxidase system together with most other components of granulocytic differentiation¹⁷. To remove constitutive transcripts, but not sequences derived from Xp21, HL60-cell-induced cDNA was enriched for Xp21 transcripts by subtractive hybridization twice with RNA from an Epstein-Barr virus (EBV)-transformed B-cell line of the

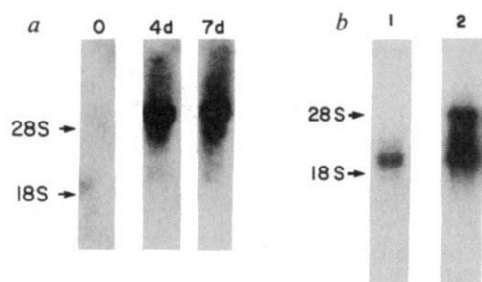


Fig. 3 Expression of 379 RNA in HL60 cells and monocytes. *a*, Total cell RNA (10 µg) isolated from uninduced (0) or dimethyl sulphoxide (DMSO)-induced (either 4 days or 7 days) human leukaemic HL60 cells was analysed with a fragment of 379 cDNA derived from the 3'-untranslated region. The uninduced cells are undifferentiated HL60 cells prepared by centrifugal elutriation¹⁷. *b*, Poly(A)⁺ mRNA (~1 µg) from human hepatoma cells (HepG2; lane 1) or from cultured human monocytes (lane 2) was similarly examined using the 379 probe plus phosphoglycerate kinase (PGK) cDNA. The latter detects a 2-kb mRNA species in all somatic cells³⁴.

Methods. RNA was prepared by guanidine-HCl precipitation and phenol extraction (*a*) or by centrifugation through CsCl₂ following solubilization with guanidine isothiocyanate³⁵ (*b*). Samples were electrophoresed in 1% agarose formaldehyde gels³⁶, transferred to nitrocellulose, and hybridized with a 1.8-kb *Bgl*II-*Bgl*II restriction enzyme fragment of H64 cDNA (see Fig. 5*a*) labelled by random priming³⁷. 28S and 18S denote the migration of ribosomal RNAs. In *b*, the filter was simultaneously hybridized with labelled insert prepared from the cDNA clone PGK-7e³⁴.

CGD/DMD patient N.F. (Fig. 1*a*). The subtracted radio-labelled cDNA, representative of ~500 individual mRNA sequences, was hybridized to a Southern blot of the Xp21 bacteriophage clones (Fig. 1*b*). Two overlapping clones, designated 379-A6 and 379-π, showed hybridization.

Isolation of cDNA clones

The results shown in Fig. 1*b* indicated that the bacteriophage clones originally isolated with pERT 379 contained sequences ultimately represented in the mRNA of induced HL60 cells. To locate the transcribed segment(s), non-repetitive DNA fragments of the 379-A6 insert were used as hybridization probes to Northern blots of the HL60 RNA used to generate the enriched cDNA. A probe consisting of three fragments (0.45, 0.5 and 1.8 kb) arising from a *Hind*III/*Bgl*II digest of a 5-kb *Hind*III fragment (designated 379-10) detected a 5-kb mRNA transcript (see below). cDNA clones corresponding to this mRNA species were isolated from 4-day-induced HL60 cDNA libraries constructed in bacteriophage λgt10 and λgt11 (refs 18–20) (Fig. 2). Repeated screening yielded cDNA clones spanning the entire mRNA transcript. Based on the frequency with which clones were obtained from the cDNA libraries, the abundance of the specific RNA in DMF-induced HL60 cells was ~0.02–0.05%. For purposes of discussion these are referred to as 379 cDNA and 379 RNA. Cloned cDNA was used to examine (1) the tissue distribution of the RNA, (2) its structure and expression in X-CGD, and (3) the nature of its predicted protein sequence.

Tissue distribution

The 379 cDNA detects a 5-kb RNA that is markedly induced in HL60 cells during granulocytic differentiation stimulated by DMF treatment (Fig. 3*a*). This RNA was also induced on treatment with either the tumour promoter TPA (12-*O*-tetradecanoyl phorbol-13-acetate) or vitamin D₃, agents that promote monocytic differentiation of HL60 cells in culture²¹. Moreover, the transcript was abundant in cultured normal human monocyte RNA (Fig. 3*b*) and absent from the hepatoma line HepG2

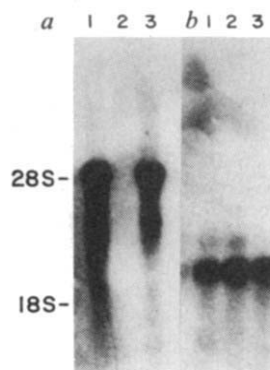


Fig. 4 Absence of 379 RNA from X-CGD monocytes. Total cell RNA (5 µg) prepared from the cultured monocytes of two X-CGD patients (lanes 1 and 2) and from a normal individual (lane 3) was examined for 379 sequences (*a*) or PGK sequences (*b*). In addition to the patient whose sample is shown in lane 2, two unrelated patients also had monocytes that were negative for 379 RNA. Peripheral blood monocytes were isolated and cultured as described previously³⁸. Other methods were as given in Fig. 3 legend.

(Fig. 3*b*) as well as from total kidney, fibroblasts and HeLa cells (not shown). A low level of the transcript was detected in normal EBV-transformed B-cell lines (not shown), which have been reported to exhibit low levels of oxidase activity²².

As the RNA transcript encoded by 379 sequences appeared to be a marker for the phagocytic lineage (granulocytes and monocytes) based on our limited survey of different cell types, it represented a suitable candidate for the product of the X-CGD locus.

Relevance of 379 transcript

Given that the X-CGD locus was mapped to ~0.1% of the genome, as many as 20 transcripts derived from this region might exist in an average cell (containing perhaps 10,000–20,000 mRNA species). This estimate does not take into account possible variation in the fraction of DNA transcribed in different regions of the genome. The finding of an Xp21 phagocyte-specific RNA cannot, by itself, be interpreted as strong evidence that it encodes the X-CGD product. To determine more specifically the potential relevance of the transcript to the disease, we analysed material from classical X-linked CGD patients by hybridization with the putative X-CGD cDNA. If the transcript were derived from the relevant disease gene locus, quantitative and/or qualitative RNA or DNA alterations would be anticipated in these genetic variants.

We first analysed RNA isolated from cultured human monocytes of both normal and CGD origin. Only CGD patients with the X-linked variety were examined. Their clinical and family histories were consistent with this classification, as were negative phagocyte NBT tests⁹ and the absence of the cytochrome *b* spectrum. Figure 4 shows the monocyte RNA analysis of two such patients: in one patient (lane 2) no 379-specific RNA was detectable; in the other patient (lane 1) 379 RNA was apparently normal in size and abundance. Integrity of RNA samples was confirmed by hybridization with a constitutively expressed sequence (phosphoglycerate kinase) derived from a different X-chromosomal region (Fig. 4*b*). Monocyte RNA of two additional X-CGD patients examined also lacked 379-specific RNA (not shown). In the RNA-negative CGD patients, Southern blot analysis with cDNA probes spanning the entire gene revealed no DNA deletions or rearrangements (not shown). These genetic variants provide strong initial support for the relevance of the 379 transcript to the disease. Although unlikely, these findings might be explained by abnormal regulation of

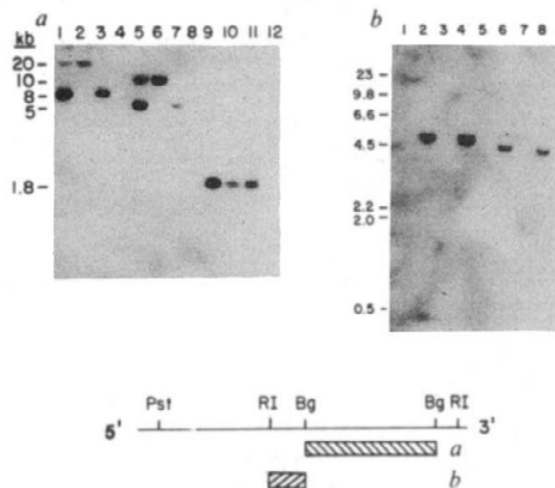


Fig. 5 Partial gene deletion in an X-CGD patient. Total cellular DNAs (5 μ g per lane) were digested with *Bam*HI (a, lanes 1–4), *Hind*III (a, 5–8; b, 1–4) or *Bgl*II (a, 9–12; b, 5–8), electrophoresed, and hybridized with the regions of the 379 cDNA indicated at the bottom of the figure. a, DNAs from mother of patient J.W. (lanes 1, 5, 9), X-CGD patient J.W. (lanes 2, 6, 10) and Xp21 deletion patient N.F.¹³ affected with CGD and DMD (lanes 4, 8, 12). Lanes 3, 7, 11, normal DNA. b, DNAs from patient J.W. (lanes 1, 5), mother of J.W. (lanes 2, 6) and patient N.F. (lanes 3, 7). Lanes 4, 8, normal DNA.

the 379 transcript due to a defect in a *trans*-acting factor produced by the authentic X-CGD locus.

Conclusive evidence regarding the role of the 379 transcript, though, was provided by an analysis of patient J.W. whose monocyte RNA appeared grossly normal in abundance and structure (Fig. 4a, lane 1). Analysis of DNA from patient J.W. with the 3' half of the cDNA revealed a DNA alteration on digestion with some (for example, *Hind*III and *Bam*HI), but not all (for example, *Bgl*II), restriction enzymes (Fig. 5a, lanes 2, 6, 10). The patient's mother was heterozygous for this alteration (lanes 1, 5), consistent with her assignment as an obligate carrier. Various subregions of the full-length cDNA were used as probes to define the nature of the DNA alteration in J.W. Use of a 0.3-kb segment encompassing the middle of the cDNA revealed an interstitial deletion in the DNA (Fig. 5b, lanes 1, 5) and monocyte RNA (not shown). Most of the region between the central *Eco*RI and *Bgl*II restriction sites of the cDNA was absent. Sequences 3' to the deleted segment were present in both the DNA and monocyte RNA (not shown). Although the precise boundaries of the interstitial deletion have yet to be defined, it overlaps the 3' terminus of the large open reading frame (ORF) of the cDNA (see Fig. 2 and below). The failure of cDNA probes to hybridize with the DNA of the CGD/DMD patient N.F. established that the transcribed sequences were derived from Xp21 (Fig. 5a, lanes 4, 8, 12).

Thus, the 379-encoded transcript was abnormal in abundance or structure in the monocytes of all four classical X-linked CGD patients examined. These observations provide conclusive evidence that the transcript defines the product of the X-CGD locus. Therefore, we refer to it henceforth as the X-CGD RNA.

Analysis of genomic DNAs of 17 additional, independent X-CGD patients with the 3' half of the cDNA revealed no detectable DNA deletions or alterations (not shown).

Sequence of X-CGD cDNA

cDNA clones spanning the 379-encoded RNA are shown in Fig. 2. The transcriptional orientation of the mRNA was estab-

lished by hybridization of single-stranded SP6 RNA transcripts²³ to Northern blots (not shown). DNA sequences of cDNA clones were obtained from both strands either from restriction enzyme fragments subcloned into M13 vectors²⁴ or by progressive deletions as described by Dale *et al.*²⁵; the assembled sequence is shown in Fig. 6a.

Although the estimated size of the RNA transcript from Northern blot analysis is nearly 5 kb, two findings indicate that the 4.27 kb of cDNA sequence that we have obtained encompasses the entire mRNA. First, the 3' terminus is defined by clone 100.1 which contained a short oligo(dA) tract. Second, primer extension analysis revealed an extended product corresponding to the most 5' extent of clone 23 (Fig. 6b). Although we cannot exclude the presence of a strong stop for reverse transcriptase, the failure to observe cDNAs with additional 5' sequences probably indicates that the true 5' terminus of the mRNA has been recognized. Genomic studies that are in progress should assign the precise transcription initiation site(s).

A single large open reading frame, extending from potential initiator ATGs at nucleotide 208 or 322 to a termination codon (TAA) at 1,726–1,728, is present. Additional ATGs, each followed closely by an in-frame termination codon, are located at positions 35 and 113. As the ATG at position 208 is in the same translational frame as that located at position 322, we cannot unambiguously assign the initiation codon. However, given that the ATG at 322 closely resembles the consensus sequence for functional initiator codons defined by Kozak²⁶, whereas that at 208 does not, it is highly likely that the former would be used for translation initiation *in vivo*. Whether the upstream ATGs have a role in regulating production of the X-CGD-encoded protein is unknown.

The 3'-untranslated region of the mRNA is 2.5 kb long. Although we have not analysed in detail the organization of the CGD gene in genomic DNA, comparison of the 379 bacteriophage subclones and cDNA clones indicated that the entire 3'-untranslated region is represented in a signal exon, which also contains a small coding segment 3' to the central *Eco*RI site of the cDNA. A putative polyadenylation signal (AATAAA) occurs 14 nucleotides before the poly(A) tract of clone 100.1. Several more typical potential polyadenylation signals (AATAAA)²⁷ are found elsewhere in the cDNA sequence but do not appear to be utilized in mRNA processing. A notable feature of the 3'-untranslated region is the presence of T₂₁ (nucleotides 4,018–4,028), followed 19 nucleotides in the 3' direction by TTTATT (nucleotides 4,057–4,062). On the complementary strand this organization [AATAAA–19 nucleotides–A₂₁] resembles the 3' end of a processed transcript^{28,29}. It is possible that a processed sequence integrated at the 3' end of the CGD gene. As the genomic DNA encoding this region has the identical sequence (not shown), this novel organization in the transcript is not the result of a cloning artefact.

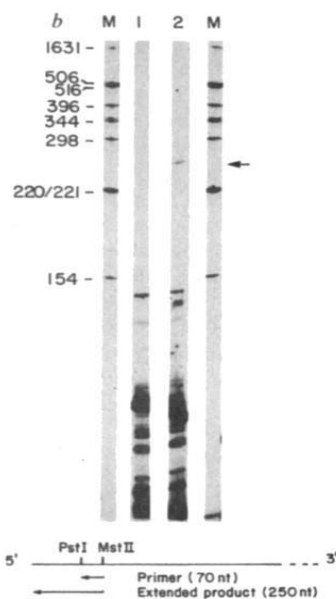
The open reading frame extending from position 322 to 1,726 predicts a primary translation product of 468 amino acids (Fig. 6) with an estimated relative molecular mass (M_r) of 54,000 (54K). If the atypical initiator ATG at position 208 were used, a polypeptide with an additional 38 amino acids at the N-terminus would result. Four potential N-linked glycosylation sites of the canonical form Asn-X-(Thr/Ser) are present. The calculated *pI* of the predicted protein is 9.5. Figure 6c shows the distribution of charged residues and a hydropathicity profile of the predicted protein. One extensive hydrophobic segment is evident; this spans amino acids 65–92. Screening of the GENBANK and protein database revealed no significant homology of the cDNA or of its predicted protein to known sequences.

Conclusions

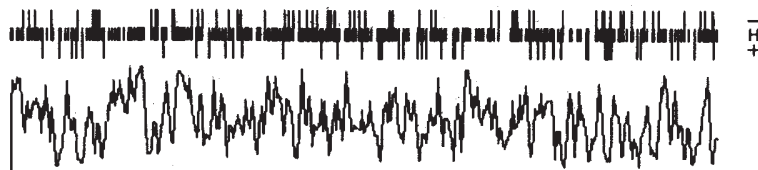
Based on genetic linkage data, chromosome walking within a defined region, and hybridization with an enriched cDNA probe, we have identified and characterized the genetic locus for an

a

CTTCTCTGCTCCACCATCGGGGAACCTGGCTGTGAATGAGGGCTCTCCATTTTCTGCTATCTGGTTGGCTGGGGTTCAACGCTCTCTCTTTGTCTGGATTACCGGGTTTATGATATT 120
 CCACCTAAGTCTTTTACACAAAGAACTCTTGGGTGAGCACTGGCACTGGCAGGGCCCTGCAGCCTGCCTGAATTTCAACTGCATCTGCTATCTTCCGCACTCTGCGAATCTG 240
 CTGCTCTCTCCAGGGGTTCCAGTGCCTGCTGCTCAACAAGAGTTCGAAGCAACCTGGACAGGAATCTACCTTTTCATATAAATGGTGGCATGGATGATTGCACTTCACTCTGGCATTCAC 360
 ACCATTGCATCTTATTAATGTGGAATGGTGTGAATCCCCGAGTCAATAATCTGATCTTATTAGTAGCACTCTCTGAACTTGGAGACAGGCAAAATGAAAGTTATCTCAATTTT 480
 GCTCGAAGAGAATAAAGAACCTGAAGGAGGCTGTACCTGGTGTGACCTGTGGAGGCATCACTGGAGTTGTCATCACGCTGTGCTCATATTAATCACTTCTCCACCAAA 600
 AlaArgLysArgIleLysAsnProGluGlyGlyLeuTyrLeuAlaValThrLeuLeuAlaGlyIleThrGlyValValIleThrLeuCysLeuIleLeuIleThrSerSerThrLys 100
 ACCATCCGGAGCTCTTACTTTCAAGTCTTTTGGTACACACATCATCTCTTCTGATCTCTTTCATTGGCCTTGCATCCATGGAGTGAACGAATTGTACCTGGGAGAGCCGAGAGAGT 720
 ThrIleArgArgSerTyrPheGluValPheTrpTyrThrHisHisLeuPheValIlePhePheIleGlyLeuAlaIleHisGlyAlaGluArgIleValArgGlyGlnThrAlaGluSer 140
 TTGGCTGTCATAATAACAGTTTGTGAACAAAAATCTCAGAATGGGGAATAAAGGAATGCCAATCCCTCAGTTTGTGGAAACCTCTTATGCTTGGAAATGGATAGTGGGT 840
 LeuAlaValHisAsnIleThrValCysGluGlnLysIleSerGluTrpGlyLysIleLysGluCysProIleProGlnPheAlaGlyAsnProProMetThrTrpLysTrpIleValGly 160
 CCCATGTTTCTGTATCTCTGTGAGAGGTTGGTGGGTTTGGCGATCTCAACAGAAGGTGGTCATCACCAGGTGGCTCACCCTTTCAAACCATCAGACTACAGATCAAGAAGAAG 960
 ArgIleValGlyAspTyrLeuCysGluArgLeuValArgPheTrpArgSerGlnGlnLysValIleThrLysValValThrHisProPheLysThrIleGluLeuGlnMetLysLys 200
 GGGTTCAAAATGGAAGTGGACAATACATTTTGTCAAGTGCCCAAGGTGTCCAAAGTGGAGTGGCACTTTTACACTGACATCCGCGCTGAGGAAGACTCTTTAGTATCCATATC 1080
 GlyPheLysMetGluValGlyGlnTyrIlePheValLysCysProLysValSerLysLeuGluTrpHisProPheThrLeuThrSerAlaProGluGluAspPheSerIleHisIle 240
 CGCATCGTTGGGACTGGACAGGGGCTGTTCAATGCTTGTGGCTGTGATAAGCAGGACTTTCAAGATGCGTGGAACTACCTAAGATAGCGGTTGATGGGCTTTGGCTGCCAGT 1200
 ArgIleValGlyAspTyrTrpGlyLeuPheAsnAlaCysGlyCysAspLysGlnGluPheGlnAspAlaTrpLysLeuProLysIleAlaValAspGlyProPheGlyThrAlaSer 300
 GAAGATGTGTCAGTATGAGTGGTGTATGTACTGGGACAGGAGTTGGGTCACACCTTCGATCCATTTCTCAAGTCAGTCTGGTCAAAATTTGCAATACGCCCACTCTGAAG 1320
 GluAspValPheSerTyrGluValMetLeuValGlyAlaGlyIleGlyValThrProPheAlaSerIleLeuLysSerValTrpTyrLysTyrCysAsnAspAlaThrAsnLeuLys 360
 CTCAAAAGATCTACTTCTACTGGCTGTGCGGGACACACATGCTTTGAGTGGTTTTCAGATCTGCTGCAACTGCTGGAGAGCCAGATGCAGAAAGGAACATGCCGGCTTCTCAGC 1440
 LeuLysLysIleTyrPheTyrTrpLeuCysArgAspThrHisAlaPheGluTrpPheAlaAspLeuLeuGlnLeuLeuGlnSerGlnMetGlnGluArgAsnAsnAlaGlyPheLeuSer 400
 TACAACTCTACCTACTGGTGGGATGAGTCTCAGGCCAATCACTTTGCTGTGACCATGATGAGGAGAAAGATGTGATCAGAGGCTGAAACAAAAGACTTTGTATGGTGGCCCAAC 1560
 TyrAsnIleTyrLeuThrGlyTrpAspGluSerGlnAlaAsnHisPheAlaValHisHisAspGluGluLysAspValIleThrGlyLeuLysGlnLysThrIleTyrGlyArgProAsn 440
 TGGGTAATGAATTCAGACAATTCAGATCAACCCCTAATACAGATAGGAGTTTCTCTGTGGACTGAGGCTTGGCTGAAACCTGAGTAAACAAAGACTCTTCAACTCTGAG 1680
 TrpAspAsnGluPheLysThrIleAlaSerGlnHisProAsnThrArgIleGlyValPheLeuCysGlyProGluAlaLeuAlaGluThrLeuLysGlnSerIleSerAsnSerGlu 480
 TCTGGCCCTCGGGGAGTGCATTTCTTCAACAAGGAACTTCAACTTGTCTCTTCCATGAGGAATAAATGTTGGGTTGTCTGCCAATGCTCAAATGCTAATGTATAATATA 1800
 SerGlyProArgGlyValHisPheIlePheAsnLysGluAsnPhe---
 AATACCCCTGCTTAAATGGACAAAAGAACTATAATGTAATGTTTTCCTTAAAGGAATGTCAAAGATTGTTTGTATGATGATAAGTTACATTTATGTGGAGCTCTATGGTTTGA 1920
 GAGCACTTTTACAAACATTTTCTTTTCTCTCAGTATGTGAGTGGAGTTTGGGAAAGATTCTGGACTCAATTTTGAATCAAAAGGGAAAGGATCAAAAGGTTGAGTAAAC 2040
 TTCCCTAAGATTATGAACCTGTGACAGATCTAGCCCATCTTACTCCAGGTTTGTATCTCTTCCACAATACTGAGCTGCCTCAGAATCTCAAAATCAGTTTATATTTCCCAAAAGA 2160
 AGAAGGAAACCAAGGATAGCTATATATTTCTACTTTGTGTCTATTTTGCATCATTTATATCATACTGAAGGAAATTTCCAGATCATTAGGACATAATACATGTTGAGAGTGCTCAA 2280
 CACTTATATGATGACAGATTTGACATCTGAGCATATCCAGTTTACTAATACAGCAGGTTAACTGGGCGAGATGTTCTTTCACAGAAGAATTTGGATTGATTGGAGTTAATGTAATCT 2400
 CATCATTTACCATCTGTCTTGGCAGAGAGCGGATCTCAAGTAAGTTTGTAAATGAATGAATGAATTTAGAACCACCAATGCCAAGATAGAATTAATTTAAAGCCTTAAACAAAAT 2520
 TATCTAAGAAATAACTTCTTACTGTCTATAGACCAAGGAATCTGATCTCCCTAGGGTCAAGAACAGGCTAAGGATATAACCAATAGGATTGCCTGAAGGGTCTGACATTTCTTA 2640
 TTTGAAGCATGAAAAAAGAGGGTTGGAGGTGGAGAATTAACCTCTGCCATGACTCTGGCTCATCTGACTGCTCTGCTCTGTCTATAAAATAAATGCAGACTAAATTTCTCTGCCAAAGTG 2760
 GTCTTCTCCAGCTAGCCCTTATGAATATTGAACCTTAGGAATTTGACAAATATGTATCTGATATGGTCAATTTGTTTAAATAACCAACCCCTTATTTTCGTAATAACACACACAAA 2880
 TGGATCGCATCTGTGACTAATGGTTTATTTGTATTATATCATCATCATCTCTAAATTAACACCCAGAAACAAAATCTCTATACAGAGATCAATTCACACTCAATAGTATGTT 3000
 CTGAATATATGTTCAAGACAGAGTCTCTAAATCACTGTTAGTGTGCCAAGAGCAGGGTTTCTTTTGTCTTAGAACTGCTCCCATTTCTGGGAATCAAAACAGTTTATTTGCCCT 3120
 ACCCTTGGAGCCCAAAATGTTTGAAGTCTTCAACTTCGGTAATGAGGAAGAGGAGAAAGCTGGGGAAGGCGAAGAGCTGGTTTGGAGGAAAGGAAATAGGAGAGAAAGAGA 3240
 ATGGGAGAGTGAGAGAAAATAAAAGGCAAAAGGAGAGAGAGGGGAAGGGGCTCTATATTTGCTCATTTCCCTGCCCAAGATTCTTAAAGTTGATATGATAGAAATATAATGAAGG 3360
 AGGTATACACATACTGATGTTGTTTGTATTCTATGTTATTTGAATCTTTTAAATCTGGTCACAAAATTTGATGCTGAGGGGGGATTTTCAAGGGACTAGGATGAATTAAGAAGCT 3480
 CAGTTGTCTTTTGTCTACTACTTCTTCTGCTCTCCCAATCTCAGGGCCTCAGGGTAGGCTGACAAATAAGGCTGCTGCTGCAATATAGCCCTTTCTGAAATGTACAGGAGTG 3600
 GTTTCTGCTTAGAGACACTTAGGTCAGCTGTTCACACTGCACCTCAGGTATCAATTTCTATTTTCAACAGATATTTATGCTGTTACTTACTGAGCTAGGCTCTGCTTTATTGTTTCAA 3720
 TTCTTTACACAAAGTATGAACGGAGGGTACCTCAGTTTATAAGGAGTCTGAGAATATGGCCCTTTCTAACCTATGTGCTAATAATTAACACAGCTTCAATTTGCTGCTCCGAGAGTG 3840
 TTCTCAAGGTTTCTATCTTCAAAACCAACTAAGTTATGAAATAGAGAGATCTGCCCTGTGTTATCCAGTTTATGAGATAAAATGAATATAAGAGTGTGCTATATAAAGTTTC 3960
 CTTTATCTCTCAAGCCACAGCTGCCAGCCACCCAGGCGAGCTGCCAGCTAGCTTTTTTTTTTTTTTTTTTTTATGCACTTATGATTTAGCAATTTATTAACAGGTACTCTAAGAAT 4080
 GATGAAGCATTTGTTTAAATCTTAAGACTATGAAGTTTCTTGTCTCTGCTTTTGGCAATGTGTTTGTGAAATTTGAATCTTGCAGGCTTGTATGTGAATATTTCTAGCGGGG 4200
 GACCTGGGAGATAATCTACCGGGGAATCTTAAACTGTGCTCAACTATAAATGAATGAGCTTCAAAAAA



c



inherited human disorder (chronic granulomatous disease) without reference to a specific protein product. The derivation of the 379 RNA transcript from Xp21, its high-level expression in the phagocytic lineage, and its derangement in patients with classical X-linked CGD provide the elements of the formal argument that the appropriate locus has been recognized.

Several aspects which contributed to the success of the transcript search are relevant to its potential application to other disorders. (1) Although only ~10% of Xp21 was represented in bacteriophage clones isolated with pERT clones¹⁵, a portion of the X-CGD gene (notably its 3'-untranslated region) was fortuitously present in the collection. (2) The 3'-untranslated region of 2.5 kb provided a large, uninterrupted exon for hybridization with the enriched cDNA probe. Although almost any transcribed segments would, in principle, be detectable by hybridization of subtracted cDNA with cloned genomic DNAs, the strength of the observed signal will be greater for larger exons. (3) Previous knowledge of cells expressing the phagocytic oxidase system suggested a convenient source for the isolation of RNA used in preparation of the enriched probe. Specifically, the abundance of the 379-encoded RNA transcript in DMF-treated HL60 cells is appreciable (on the order of 0.02–0.05%). (4) The availability of an EBV-transformed cell line from patient N.F. with a deletion through the entire Xp21 region permitted removal of most constitutive transcripts from the enriched cDNA without depletion of Xp21 sequences. The extent to which our strategy may be applied to the isolation of transcripts for other disorders for which a chromosomal map position is established may well depend on these factors.

Although the oxidase system of phagocytes has been studied by conventional biochemical analyses for two decades, no clear view of the components involved or the specific protein abnormalities manifest in the major functional deficiency seen in CGD has emerged. Recently, the findings of Segal and colleagues

have implicated an unusual *b*-type cytochrome in the oxidase system^{11,30,31}. As such, it represented a plausible candidate for the protein mutated in the X-linked form of CGD^{11,30,31}. Our results, however, argue against this conclusion. First, the amino-acid composition of the protein predicted by our sequence does not resemble that reported for purified cytochrome *b*³⁰. Second, the predicted protein encoded by the 379 RNA does not show significant homology to previously sequenced cytochromes of several origins. No evidence for a haem-binding region was evident from analysis of the predicted protein. Third, although its spectrum is absent in CGD granulocytes, the proposed cytochrome *b* protein has been reported to be present in extracts of neutrophils from affected individuals³¹. This finding is not in accord with the apparently frequent observation of RNA-negative X-CGD patients. Whether the predicted protein is related to the flavoprotein found in phagocytes¹² must await further biochemical characterization of the latter.

We propose that the 379-encoded protein, which we designate the X-CGD protein, is an essential component of the oxidase system of the phagocyte and probably interacts with other proteins to form an active complex. Among those components that may interact with the X-CGD protein are the membrane-associated cytochrome *b* and a flavoprotein. Spectral abnormalities of cytochrome *b* in CGD phagocytes are probably secondary to the absence (or abnormal structure) of the predicted X-CGD protein.

Characterization of the X-CGD protein *in vivo* and examination of proteins with which it associates should help to elucidate the organization and function of the oxidase system of the phagocyte. Analysis of the predicted sequence of the X-CGD protein has provided few clues, as yet, to the nature of the protein itself. The protein has an estimated *M_r* of a minimum 54K, and is probably basic. Modification at the four potential N-linked glycosylation sites may alter the apparent size of the observed protein *in vivo*. One marked hydrophobic region noted above (amino acids 65–92) may represent a transmembrane segment. Finally, the interstitial deletion found in patient J.W. may serve to locate an important functional domain at the carboxy-terminus of the protein, as the deletion is predicted to remove the last 50 amino acids (distal to the *EcoRI* site at nucleotide 1,571). Further characterization of the X-CGD protein and delineation of its functional domains will ultimately require the introduction and expression of the cDNA in phagocytic cells of X-CGD patients.

We thank Dr Temple Smith of the Dana-Farber Cancer Institute for assistance and advice in computer analysis of the sequence and Dr Mark Davis for advice on subtractive hybridization. S.H.O. is an Investigator of the Howard Hughes Medical Institute. P.E.N. was supported by NIH grant CA38325, by the Socha Memorial Fund for CGD research, and by an Established Investigatorship of the American Heart Association with funds contributed in part by its Massachusetts affiliate. This work was also supported in part by grants from the Muscular Dystrophy Association to L.M.K. and from the NIH to J.T.C. (AI-21320) and S.H.O. (HD18661).

Received 17 May; accepted 21 May 1986.

Fig. 6 (Left) DNA sequence of the X-CGD cDNA (a), primer extension analysis (b) and charge distribution and hydropathy profile of predicted X-CGD protein (c). **a**, Bacteriophage cDNA inserts were cloned in M13 derivatives²⁴ for sequence analysis by the methods of Sanger *et al.*³⁹ and Dale *et al.*²⁵. All regions were sequenced on both DNA strands. Clones from which sequences were derived are shown in Fig. 2. H64 and H61 were sequenced in their entirety. The initial 0.1-kb displayed was assembled from clones 23, 101, 403 and 99. The 2 kb spanning the open reading frame was derived from clone 99. Overlaps with the 5' regions of H64, H61 and 405 were examined. Based on this, and restriction mapping of genomic DNA, the intron sequences at the 5' end of H61 were identified. The extreme 3'-terminal sequences were obtained from clone 100.1. The first four ATGs in the sequence are underlined, with the predicted initiator codon double-underlined. The presumed processed cDNA sequence near the 3' end of the transcript is boxed. Canonical N-glycosylation sites are underscored with broken lines, and the presumptive poly(A) addition signal (ATTAAG) is overlined. **b**, The position of the 5' terminus of the X-CGD mRNA was assigned by primer extension analysis. A 70-nucleotide (nt) end-labelled primer was hybridized with yeast transfer RNA (10 µg, lane 1) or total monocyte (20 µg, lane 2) and extended with reverse transcriptase. Lanes labelled M, marker DNA. The observed product of 250 nt corresponds precisely to the 5' extent of clone 23 (see Fig. 2). Use of a synthetic 30-mer primer (nucleotides 80–110) confirmed the position of the extended product and suggested, in addition, the existence of three possible mRNA start sites within about 10 nucleotides (not shown). **c**, Distribution of charged residues and the hydropathy profile of the predicted X-CGD protein. +, H and – refer to cationic (Arg, His, Lys), hydrophobic (Ala, Ile, Met, Trp, Val, Phe, Pro, Leu) and anionic (Glu, Ser, Tyr, Asp, Thr) amino acids, respectively. In the hydrophobicity profile, shown at the bottom, upward deflections indicate hydrophobic regions. **Methods.** **a**, The indicated 70-nt *MstII*–*PstI* restriction fragment was 5'-end-labelled with [γ -³²P]ATP using polynucleotide kinase⁴⁰, hybridized with RNA, and extended with AMV reverse transcriptase using methods described previously⁴¹. Reaction products were phenol-extracted, ethanol-precipitated and separated in an 8% urea/acrylamide gel. Marker DNAs were pBR322 digested with *HinfI*. Autoradiography was performed overnight with an intensifying screen.

1. Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. *Am. J. hum. Genet.* **32**, 314–331 (1980).
2. Davies, K. E. *et al. Nucleic Acids Res.* **11**, 2303–2312 (1983).
3. Monaco, A. P. *et al. Nature* **316**, 842–845 (1985).
4. Knowlton, R. G. *et al. Nature* **318**, 380–382 (1985).
5. White, R. *et al. Nature* **318**, 382–384 (1985).
6. Wainwright, B. J. *et al. Nature* **318**, 384–385 (1985).
7. Gusella, J. F. *et al. Nature* **306**, 234–238 (1983).
8. Collins, F. & Weissman, S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6812–6818 (1984).
9. Tauber, A. I., Borregaard, N., Simons, E. & Wright, J. *Medicine* **62**, 286–309 (1983).
10. Weening, R. S. *et al. J. clin. Invest.* **75**, 915–920 (1985).
11. Segal, A. W. *et al. New Engl. J. Med.* **308**, 245–251 (1983).
12. Markert, M., Glass, G. A. & Babior, B. M. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3144–3148 (1985).
13. Baehner, R. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 3398–3401 (1986).
14. Francke, U. *et al. Am. J. hum. Genet.* **37**, 250–267 (1985).
15. Kunkel, L. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 4778–4782 (1985).
16. Collins, S. J., Ruscetti, F. W., Gallagher, R. E. & Gallo, R. C. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2458–2462 (1978).
17. Newburger, P. E. *et al. J. biol. Chem.* **259**, 3771–3776 (1984).
18. Young, R. A. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1194–1198 (1983).

19. Gubler, U. & Hoffman, B. J. *Gene* **25**, 263-279 (1983).
20. Ginsburg, D. *et al. Science* **228**, 1401-1406 (1985).
21. Rovera, G., Santoli, D. & Damsky, C. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2779-2783 (1979).
22. Volkman, D. J., Buescher, E. S., Gallin, J. I. & Fauci, A. S. *J. Immunol.* **133**, 3006-3009 (1984).
23. Melton, D. A. *et al. Nucleic Acids Res.* **12**, 7035-7056 (1984).
24. Messing, J. *Meth. Enzym.* **101**, 20-78 (1983).
25. Dale, R. M. K., McClure, B. A. & Houchins, J. P. *Plasmid* **13**, 31-40 (1985).
26. Kozak, M. *Nucleic Acids Res.* **12**, 857-872 (1984).
27. Proudfoot, N. J. & Brownless, G. G. *Nature* **263**, 211-214 (1976).
28. Nishioka, Y., Leder, A. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2806-2809 (1980).
29. Sharp, P. A. *Nature* **301**, 471-472 (1983).
30. Harper, A. M., Chaplin, M. F. & Segal, A. W. *Biochem. J.* **227**, 783-788 (1985).
31. Segal, A. W. *Lancet* **i**, 1378-1382 (1985).
32. Davis, M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2194-2198 (1984).
33. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
34. Michelson, A. M., Markham, A. F. & Orkin, S. H. *Proc. natn. Acad. Sci. U.S.A.* **80**, 472-476 (1983).
35. Chirwin, J. M., Przygla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5298 (1979).
36. Maniatis, T., Fritsch, E. & Sambrook, J. *Molecular Cloning: a Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
37. Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **132**, 6-13 (1983).
38. Strunk, R. C., Whitehead, A. S. & Cole, F. S. *J. clin. Invest.* **76**, 985-990 (1985).
39. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
40. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-580 (1980).
41. Treisman, R., Proudfoot, N. J., Shander, M. & Maniatis, T. *Cell* **29**, 903-911 (1982).
42. Francke, U. *Cytogenet. Cell Genet.* **38**, 298-307 (1984).

LETTERS TO NATURE

A QSO with redshift 3.8 found on a UK Schmidt telescope IIIa-F prism plate

C. Hazard*†, R. G. McMahon*† & W. L. W. Sargent‡

* Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

† Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

‡ Palomar Observatory, California Institute of Technology, Pasadena, California 91125, USA

High-redshift QSOs (quasi-stellar objects) are important because of the information they provide on the early history and evolution of the Universe; the brighter ones are particularly valuable as probes of the intervening material. During the past 20 years much effort has been devoted to trying to discover QSOs at redshifts $z > 3.5$, but progress has been so slow that it has been suggested that there is a cutoff in the QSO distribution at $z \approx 3.5$ (ref. 1). However, we have already demonstrated²⁻⁴ how unfiltered UK Schmidt telescope (UKST) IIIa-F low-dispersion objective prism plates can be used successfully in such searches up to at least $z \approx 3.7$ (ref. 3). We report here the use of this technique to discover a QSO, 1208 + 1011, with a redshift $z = 3.80$. The highest redshift previously known was that of the radio-selected QSO, PKS2000-330 with $z = 3.78$ (ref. 5). In two UKST fields we have now discovered six QSOs with redshifts between 3.3 and 3.8, of which four have $z \geq 3.50$; including PKS2000-330, only four other QSOs with $z \geq 3.50$ are known over the whole sky^{5-7,10}. Our success up to $z = 3.8$ indicates that redshifts > 4 could soon be attained.

Up to $z \approx 3.3$, searches for high- z QSOs by their Ly- α + N v emission are best carried out on IIIa-J prism plates, which have a red cutoff to their spectral response at $\sim 5,200 \text{ \AA}$. The IIIa-F plates, with a corresponding limit at $6,800 \text{ \AA}$, extend the range above $z \approx 3.3$ to a possible maximum of $z = 4.6$. However, to avoid confusion with cool stars our searches on these plates have been restricted to the interval $3.3 \leq z \leq 3.9$. With this restriction and using auxiliary IIIa-D and IIIa-J prism and direct J and R or I plates to check the spectral features and colours of the IIIa-F-selected objects we have shown how high-redshift QSOs can be selected with complete reliability^{2,3}. For these prime selections we require that the reality of the emission line in the IIIa-F spectrum be confirmed by the presence of a corresponding emission line in the IIIa-D spectrum and that its identification with Ly- α + N v be confirmed by the detection of O VI emission on the IIIa-J spectrum. We also look for a fall in the continuum level going from red to blue across Ly- α + N v and check for, but do not demand, a cutoff near the Lyman limit. When these conditions are met, the selections are no longer candidate high-redshift QSOs but certain QSOs with redshifts within about ± 0.05 of those measured from the prism spectra; the accurate redshifts now available for our other high- z selections have allowed us to correct a slight systematic bias of $+0.03$ in our earlier prism redshifts. These prime selections are necessarily restricted to bright QSOs with strong O VI as well

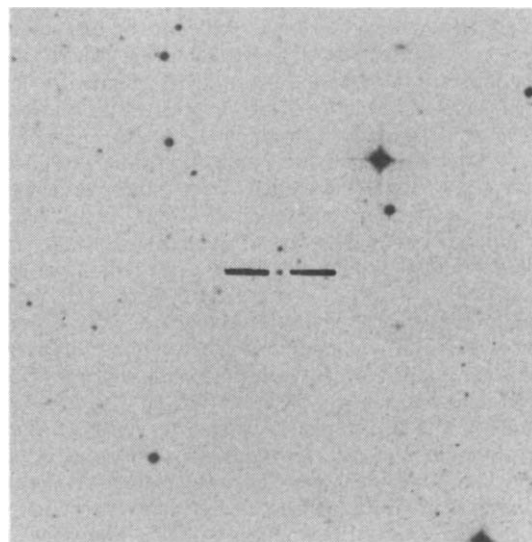


Fig. 1 Finding chart of 1208 + 1011, 8×8 arc min reproduced from a direct IIIa-J UKST plate (courtesy of the UKST unit). North is up, and east is to the left.

as strong Ly- α + N v emission and including 1208 + 1011 we have so far found only five objects²⁻⁴ which could confidently be classified as QSOs with $z \geq 3.3$ from the prism data alone; a sixth, the broad absorption line (BAL) QSO 0105-2634 with $z = 3.50$, appeared a certain QSO from the prism data but was relegated to the 'candidate' category because of its unusual colours⁴. As among the numerous IIIa-F selected candidates there may be high- z QSOs with weak O VI, and because we are still refining our selection criteria, we have not confined our follow-up slit spectroscopy to the certain selections. However,

Table 1 Measured wavelengths of emission lines, with corresponding redshifts

Emulsion	Feature	Wavelength* ($\text{\AA}$)	Identification	Redshift*
IIIa-F	Emission	$5,860 \pm 50$	Ly- α + N v	3.82 ± 0.04
IIa-D	Emission	$\sim 5,950^\dagger$	Ly- α + N v	~ 3.89
IIIa-J‡	Emission	$4,980 \pm 80$	O VI	3.82 ± 0.08
Adopted redshift				3.82 ± 0.04

* The errors are maximum errors. The calibration curve for the IIIa-F spectra was normalized using Ly- α + N v in QSOs of known redshift, assuming the wavelength of the blend to be $1,216 \text{ \AA}$. It automatically takes into account the N v contribution unless it differs markedly from the average.

† Ly- α + N v is too close to the red limit of the IIa-D spectrum to be accurately measured. The quoted wavelength is based on measurements relative to nearby stars.

‡ The IIIa-J spectrum dispersion is perpendicular to that of the IIIa-F spectrum, and is thus not confused with the nearby star.

Table 2 Seven QSOs with $z \geq 3.3$ visible on the 0053–2803 and 1204+1129 IIIa-F prism plates

Name*	Position (Epoch 1950.0)		$m_R^\dagger$	$M_R^\ddagger$	Equivalent width Ly- α + N v	Z_{em}
	RA	Dec.				
0042–2627	00 h 42 min 06.22 s	–26° 27' 42.9"	17.0	–29.3	50	3.30§
DHM0054–284	00 53 59.8	–28 24 45	17.8	–28.7	60	3.61
0055–2659	00 55 32.46	–25 29 26.0	17.1	–29.5	15	3.67
0105–2634	01 05 48.19	–26 34 20.2	17.3	–29.2	50 (BAL)	3.50
1159+1223	11 59 14.23	+12 23 11.9	17.5	–29.0	150	3.51
1208+1011	12 08 23.73	+10 11 07.9	17.5	–29.2	46	3.80
1209+0919	12 09 01.66	+09 19 04.0	18.5	–27.8	120	3.31

* All except the colour-selected DHM0056–284⁷ were discovered in the present programme.

[†] The magnitude of the BAL 0105–2636 has been corrected for heavy absorption in the R pass-band. Magnitudes are accurate to ± 0.5 mag.

[‡] Calculated for $q_0 = 0.5$, $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and assuming a continuum spectra index of -0.5 .

§ This is lower than the previously quoted $z = 3.4$, which was a provisional estimate made at the telescope.

of ~ 20 'candidates' all but the BAL QSO 0105–2634 have turned out to be galactic stars.

Until recently we possessed only one set of high-quality prism plates, centred on (RA, dec.) = (00 h 53 min, $-28^\circ 03'$), and on which we found³ the $z = 3.67$ QSO 0055–2659. The $z = 3.51$ QSO 1159+1223 had been found earlier² on a set of lower-quality plates, centred on (12 h 04 min, $11^\circ 29'$), but it is an exceptional object with the observed equivalent width of Ly- α + N v = 700 Å. The QSO 1208+1011 was found on a new high-quality IIIa-F plate of this latter region taken at an exposure of 20 min and with a limiting magnitude $m_R = 19$. On it, 1159+1223 is much more prominent than on the earlier discovery plate on which 1208+1011 is not even recognizable as a QSO. The IIIa-F spectra of both 1159+1223 and 1208+1011 from the new plate, which are reproduced in Fig. 1, with Ly- α + N v indicated in each case, show clearly the higher redshift of 1208+1011. The sharp drop in the continuum level, going from red to blue across Ly- α + N v and which is particularly prominent in 1208+1011, is a characteristic of high- z QSOs and in itself a valuable selection criterion. The lower visibility of the emission line in 1208+1011 shows that it has a significantly smaller observed equivalent width than the 700 Å of the emission line in 1159+1223. The emission line is also visible on the Ila-D spectrum, but too close to the plate's red limit to permit an accurate wavelength measurement. Its identification with Ly- α + N v was confirmed by the IIIa-J spectrum, which shows a second emission line at the expected position of O VI. This IIIa-J spectrum, with dispersion perpendicular to that of the IIIa-F spectrum, and therefore not confused by the nearby star, also shows a highly unusual continuum distribution, with a steep decrease in level at $\sim 4,250$ Å, below which wavelength it continues faintly before disappearing at $\sim 3,700$ Å. The decrease in continuum level is ~ 150 Å to the blue of the Lyman limit at the emission line redshift, but relatively few QSOs show strong absorption at the Lyman limit.

Table 1 lists the measured wavelengths of the emission lines and the corresponding redshifts. The calibration curve used for the IIIa-F plate was normalized using the Ly- α + N v emission in other QSOs of known redshift. The wavelength of the Ly- α + N v blend was assumed to be 1,216 Å, but the method of calibration automatically takes into account the N v contribution unless it differs markedly from the average. The redshifts calculated from Ly- α + N v and O VI are in such excellent agreement as to remove any doubt about the line identifications or our adopted redshift of 3.82 ± 0.04 . These quoted errors are maximum errors, giving a minimum possible redshift for 1208+1011 of 3.78, equal to that of PKS2000–330⁵. However, a careful comparison of the IIIa-F spectrum of 1208+1011 with those of 1159+1223, DHM0054–284 and 0059–2659, all with redshifts > 3.5 (refs 2, 3, 7), showed that a redshift as low as 3.78 was highly improbable. The QSO 1208+1011 was therefore selected as our highest priority object for further study, and a slit spec-

trum was obtained using the 5-m Hale telescope on Mount Palomar, which gives an accurate redshift of $z = 3.803 \pm 0.005$ (see ref. 8).

1208+1011 is the sixth QSO with $z > 3.3$ to be selected from the 0053–2803 and 1204+1129 low-dispersion IIIa-F prism plates and for which slit spectra are now available. The data on all six objects are summarized in Table 2, together with those for the $z = 3.61$ QSO DHM0054–284⁷, which is also recognizable on the 0053–2803 plate. Their rest-frame equivalent widths, W_λ , of Ly- α + N v are consistent with the spread of W_λ in a sample⁹ of IIIa-J-selected QSOs having $2.6 < z < 3.3$, with four lying towards the lower limit ($W_\lambda = 50$) of this distribution. This lower limit is not an inherent observational limit but reflects only that the preliminary surveys with both IIIa-J and IIIa-F plates concentrated on strong-emission-line objects. Other than the objects in Table 2, only eight QSOs with $z > 3.3$ are known^{10,11}. If attention is restricted to the higher-redshift objects, the potential of the IIIa-F surveys becomes even more striking. It lists five QSOs with $z \geq 3.50$ visible in two UKST fields with a useful area of $\sim 60 \text{ deg}^2$. There are only three other QSOs known above this redshift over the whole sky.

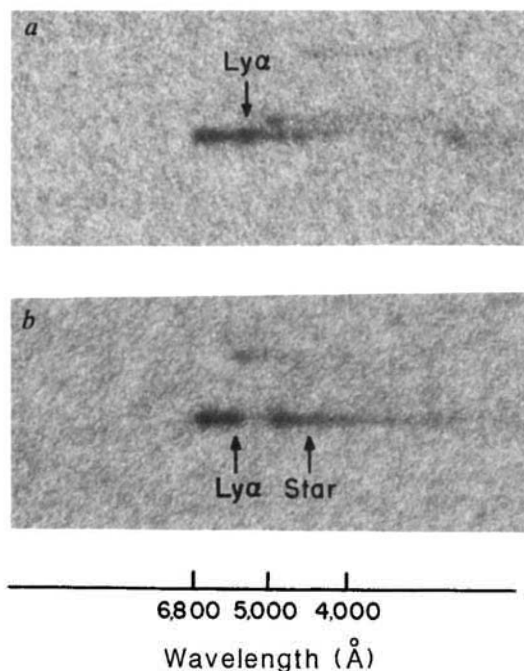


Fig. 2 IIIa-F prism spectra of the $z = 3.51$ QSO 1159+1223 (a) and of 1208+1011 (b) showing clearly the higher redshift ($z = 3.80$) of the latter object. The spectrum to the immediate right of 1208+1011 is that of a galactic star. (Courtesy of the UKST unit.)

The significance of 1208+1011 is not so much its high redshift but that it is the third QSO with $z > 3.3$ to be found on the 1204+1129 plate. This effectively eliminates the possibility that the high density of such QSOs on the 0053–2802 plate³ represented only a violent statistical fluctuation or was indicative of QSO clustering. With the results from the two plates we have now established not only that QSOs with $z > 3.5$ exist in relatively large numbers but that they may be found on unfiltered IIIa-F objective prism plates at a rate of $\sim 0.1 \text{ deg}^{-2}$ for $3.5 < z < 3.8$ and $m_R \leq 18$. A remarkable feature is that this relatively high density is for objects with absolute magnitudes $M_R \leq -28.5$ (calculated for deceleration parameter $q_0 = 0.5$, Hubble constant $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and a continuum spectral index of -0.5). At $z = 2.25$ similar QSOs would have apparent magnitudes $B \leq 17.5$, but the observed surface density for $B \leq 17.5$ in the redshift interval $2.0 < z < 2.5$ is only 0.01 deg^{-2} (ref. 12). The interval $2.0 < z < 2.5$ corresponds to a co-moving volume twice that of the interval $3.5 < z < 3.8$, so that, even allowing for uncertainties in the magnitude estimates, our incomplete observations imply a strong increase in the co-moving density of the brightest QSOs with increasing redshift. Our success in finding bright QSOs with $z > 3.5$ contrasts with the reported null detections in deep surveys over areas of $1\text{--}5 \text{ deg}^2$ for QSOs with $3.5 < z < 4.7$ and $m_R \leq 21$ (refs 1, 13, 14). The results of these deep surveys appear incompatible with our observations unless the increasing density of QSOs with $M_R < -28.5$ is accompanied by a decrease in the co-moving density of lower-luminosity QSOs with $-28.5 < M_R < -26$, the lower limit of $M_R < -26$ corresponding to $m_R \leq 21$ at $z = 3.5$. This implies a transition from the steep QSO luminosity function found at lower redshifts¹² to one that is relatively flat above $z = 3.5$ or perhaps even one where the number of QSOs per magnitude interval decreases at lower luminosities. For an unchanged flat luminosity function and constant co-moving density above $z = 3.5$ our observations would imply 1 QSO deg^{-2} with $m_R < 21$ in the redshift range $3.5 < z < 4.7$.

Above $z \approx 4.2$, Ly- α + N V will lie so close to the red limits of the low-dispersion IIIa-F prism spectra that it will be very difficult to distinguish it from the enhancement in the continuum level produced by the decreased prism dispersion which is a feature of most low-dispersion IIIa-F prism spectra and particularly those of cool stars. However, there now appears to be no problem in detecting it up to $z \approx 4.2$ because we have already seen C IV at $\lambda = 6,200 \text{ \AA}$ in QSOs with $z \approx 3$. The problem in redshift range is that above $z \approx 3.9$, Ly- α passes beyond the red limit of the IIa-D emulsion and O VI moves close to the red limit of the IIIa-J emulsion so that the auxiliary plates no longer provide a useful check on the selections. Typically, our searches yield about 30 candidates per plate with $3.9 < z < 4.2$. Assuming that the properties and surface density of QSOs in the redshift range do not differ significantly from those in the range $3.5 < z < 3.9$, we would expect only one or two of these to be genuine QSOs. About 20 candidates will need to be studied spectroscopically to give a reasonable chance of finding one QSO with $z > 3.9$. However, with modern instrumentation this is not a particularly formidable task, requiring about one night's observing. Also, our experience with IIIa-J prism plates has shown that the results of follow-up slit spectroscopy soon allows a refinement of the selection procedures so that the number of candidates per plate should be reduced significantly. The discovery of QSOs with $z > 4$ would now appear to require only an adequate number of high-quality UKST prism plates and time to follow up the candidates on large telescopes. As we have already noted³, for rare objects with a flat luminosity function, the large area which can be covered by Schmidt plates in such surveys is a more significant consideration than their relatively bright limiting magnitude. Our results so far have demonstrated that these surveys can be more productive than deep surveys over small areas of sky.

We thank UKST for their help and plate material, without which our studies would not have been possible.

Received 27 February; accepted 29 April 1986.

1. Osmer, P. S. *Astrophys. J.* **253**, 28–37 (1982).
2. Hazard, C. *et al. Mon. Not. R. astr. Soc.* **211**, 45P–50P (1984).
3. Hazard, C. & McMahon, R. G. *Nature* **314**, 238–240 (1985).
4. Hazard, C., McMahon, R. G. & Morton, D. C. *Mon. Not. R. astr. Soc.* (submitted).
5. Peterson, B., Savage, L. A., Jauncey, D. L. & Wright, A. E. *Astrophys. J.* **260**, L27–L29 (1983).
6. Wampler, E. J., Robinson, I. B., Baldwin, J. A. & Burbidge, E. M. *Nature* **243**, 336–337 (1973).
7. Shanks, T., Fong, R. & Boyle, B. J. *Nature* **303**, 156–158 (1983).
8. Sargent, W. L. W., Filippenko, A. V., Steidel, C., Hazard, C. & McMahon, R. G. *Nature* **322**, 40–42 (1986).
9. Hazard, C., Morton, D. C., McMahon, R. C., Sargent, W. L. W. & Terlevich, R. *Mon. Not. R. astr. Soc.* (in the press).
10. Dunlop, J. S. *et al. Nature* **319**, 564–567 (1986).
11. Veron-Cetty, M.-P. & Veron, P. *A Catalogue of Quasars and Active Nuclei* (Eur. Southern Observ. scient. Rep. No. 4, 1985).
12. Weedman, D. *Astrophys. J. Suppl. Ser.* **57**, 532 (1985).
13. Koo, D. C. in *Proc. 24th Liège. int. Astrophys. Colloq.* (ed. Swings, J. P.) (University of Liège, 1984).
14. Schneider, D., Schmidt, M. & Gunn, J. *Bull. Am. astr. Soc.* **16**, 488 (1984).

Spectrum of a QSO with redshift 3.8

Wallace L. W. Sargent*, Alexei V. Filippenko†, Charles C. Steidel*, Cyril Hazard‡ & Richard G. McMahon§§

* Palomar Observatory, California Institute of Technology, Pasadena, California 91125, USA

† Department of Astronomy, University of California, Berkeley, California 94720, USA

‡ Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

§ Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

Hazard *et al.*¹ have described the circumstances leading to the identification of a quasi-stellar object (QSO) (1208+1011) with $z = 3.82 \pm 0.04$. Here we present a moderate-resolution optical spectrum of this object. Its precise redshift, $z = 3.803 \pm 0.005$, is the highest ever recorded.

The data were obtained on 13 January 1986 UT with the double spectrograph² at the Cassegrain focus of the 5.08-m Hale telescope at Palomar Observatory. A slit of width 2 arc s was aligned along the parallactic angle corresponding to the approximate midpoint of the 4,000-s exposure. Observations were made through clouds whose extinction was visually estimated to be ~ 1 mag. The average air mass (density) was ~ 1.11 . A dichroic filter directly behind the slit reflected blue light ($\lambda \leq 5,500 \text{ \AA}$) to one camera and transmitted red light to the other. The detector on the blue side was the '2D-Fruti' photon counting system devised by Shectman³. When used with a grating with 600 grooves mm^{-1} and blazed at $3,780 \text{ \AA}$ in first order, this arrangement gives spectra covering the range $3,200\text{--}5,600 \text{ \AA}$ at a resolution of $\sim 6 \text{ \AA}$. The detector on the red side of the spectrograph was a Texas Instruments 800×800 CCD (charge-coupled device). A grating with 158 grooves mm^{-1} blazed at $7,560 \text{ \AA}$ in first order produced a resolution of $\sim 18 \text{ \AA}$ over the range $5,400\text{--}10,200 \text{ \AA}$.

Calibrations were performed in the usual manner⁴, and fluxes were derived with the standard stars^{5,6} GD–248 (for the blue spectrum) and BD+26°2606 (for the red). Some uncertainty is present in the continuum shape of the QSO between $\sim 5,100$ and $5,500 \text{ \AA}$ due to the low signal-to-noise ratio in this region of the spectrum of GD–248. The two individual spectra were combined to yield one complete scan over the wavelength range $3,200\text{--}10,200 \text{ \AA}$, as shown in Fig. 1. Although clouds affected the QSO observations, data for BD+26°2606 were obtained in similar conditions, and the absolute flux scale is probably accurate to within $\sim \pm 0.6 \text{ mag}$ (1σ).

Figure 1 confirms that 1208+1011 is indeed a QSO of high redshift. Prominent, broad emission lines of Lyman- β + O VI, Ly- α + N V, Si IV + O IV], C IV, and Al III + C III] are present.

Table 1 Measurements of emission lines, 1208+1011

Line, λ_{vac} (Å)	Observed λ_{air} of peak (Å)	z (peak)*	$R(\lambda)/F(C\text{ IV})^\dagger$	Observed W (Å)‡
Ly- β , 1,025.7	—	—	0.46	53
O VI, 1,033.8	4,971	3.810		
Ly- α , 1,215.7	5,873	3.832		
N V, 1,240.1	5,946	3.796	2.29	220
Si II, 1,260.4	6,090	3.83		
O I, 1,303.5	6,300	3.82	0.13	12
Si II, 1,307.6	6,300	3.82	0.07	6
? 1,342§	6,442	—	0.08	7
Si IV, 1,396.8	6,732	3.804	0.28	27
O IV], 1,402.5	7,437	3.802	1.00	107
C IV], 1,549.1	7,850	3.79		
He II, 1,640.4	8,950	3.82	0.27	<5¶
Al III, 1,858 #	9,170	3.806	0.35	35
C III], 1,908.7	9,990	—	0.36	47
? 2,080§	—	—	—	50

* Assumptions: Ly- β negligible compared to O VI; $\lambda(O\text{ I} + \text{Si II}) = 1,306$ Å; $\lambda(\text{Si IV} + \text{O IV}) = 1,401.6$ Å; Si III] 1,892 Å negligible compared to Al III 1,858 Å.

† Flux of C IV 1,549 Å = 1.77×10^{-14} erg s $^{-1}$ cm $^{-2}$. Galactic extinction $A_V \approx 0.0$ mag.

‡ Observed equivalent width. Divide by $1+z=4.803$ to get W in QSO rest frame.

§ Derived from $z=3.803$ and the observed λ ; exact identity unknown.

|| Very approximate wavelength and redshift.

¶ Upper limit; line not detected with certainty.

Possibly contaminated by Si III] 1,892 Å and Fe II 1,860 Å.

The velocity widths of Ly- α and C IV are $\sim 6,000$ km s $^{-1}$ (full width at half-maximum) and $\sim 18,500$ km s $^{-1}$ (full width at zero-intensity). There is a sharp decrease in the continuum flux blueward of the Ly- α emission. This is undoubtedly due to the Ly- α absorption-line 'forest' which is so typical of distant QSOs⁷⁻⁹, and the spectrum consequently appears quite red.

Table 1 lists the observed emission lines, together with their strengths, equivalent widths, wavelengths, and redshifts. The relative intensities of strong lines are uncertain by $\sim \pm 15\%$, primarily because of difficulties in defining the continuum. Weak lines obviously have larger errors. Regions containing several closely-spaced emission lines (for example, that between Ly- α and Si IV+O IV]) exhibit a deceptively high continuum due to the blending of the line wings, so the quoted fluxes are lower limits. Extinction produced by our Galaxy is negligible at the latitude of 1208+1011 ($b \approx 70^\circ$).

A redshift of $z=3.803 \pm 0.005$ is derived from a weighted average of the apparent peaks of C IV 1,549.1 Å, {Si IV+O IV} 1,401.6 Å (ref. 10) and C III] 1,908.7 Å. Ly- α is not used in the average, because its blue side is eroded by numerous absorption lines. O VI is also severely affected. Although Table 1 lists several weak emission lines, they are excluded from the average because of large uncertainties in their measured wavelengths.

If redshifts are of cosmological origin, as is generally assumed, 1208+1011 is now the most distant known object in the Universe; PKS2000-330, the previous champion¹¹, has $z=3.78 \pm 0.01$. Like many other high-redshift QSOs, 1208+1011 is exceedingly luminous: we find $m(6,960 \text{ Å}) \approx 17.8$, comparable to that of PKS2000-330. If the Hubble constant $H_0 = 50$ km s $^{-1}$ Mpc $^{-1}$, this corresponds to an absolute magnitude $M \approx -30.8$ (for deceleration parameter $q_0 = 0$) or $M \approx -28.5$ (for $q_0 = 1$) at a rest wavelength of 1,450 Å.

With the exception of He II 1,640 Å, which is unusually weak and may even be absent, the relative strengths of emission lines in 1208+1011 are similar to those observed in many QSOs¹². Although the Ly- α /C IV ratio is slightly lower than normal, the observed flux of Ly- α is decreased by the forest of absorption lines on the blue side of the profile. High-resolution spectra are needed to quantify the effect of these lines. The C III]/C IV ratio is also fairly small, but part of the C III] flux may have been incorrectly attributed to Al III 1,858 Å. The equivalent width (W_0) of C IV 1,549 Å in the QSO's rest frame is 22.3 Å, lower than the average value (35 ± 25 Å) found by Baldwin¹² in a sample of high-redshift QSOs. Given the uncertainty in the measured magnitude of 1208+1011, however, this W_0 (C IV) is quite consistent with the correlation he discovered¹³ between W_0 (C IV) and the continuum luminosity at $\lambda = (1+z)1,450 \text{ Å}$: $\log L_{\text{cont}} = 33.34 - 1.57 \log W_0$ (for $q_0 = 1$).

Note that a broad emission line appears to be present at a rest wavelength of $\sim 2,080$ Å. This feature has previously been seen in the spectra of some QSOs¹⁴⁻¹⁶, but its exact identity remains unknown. One possibility is permitted Fe II, as suggested by Arp¹⁵.

The continuum in the range 3,200-5,600 Å undulates in a peculiar manner. It is possible that the strong feature around 4,050 Å represents the Lyman limit of absorbing clouds at $z \approx 3.44$, but the deep, rather broad dips near 5,500 Å, 4,560 Å, and 4,410 Å are more consistent with Ly- α , Ly- β , and Ly- γ absorption (respectively) at $z \approx 3.53$. Better spectra are needed to confirm this, as well as to investigate the properties of the Ly- α forest. Unlike the case of PKS2000-330, continuum radiation is visible all the way down to the atmospheric ozone cutoff at $\sim 3,200$ Å. This precludes the presence of very optically thick

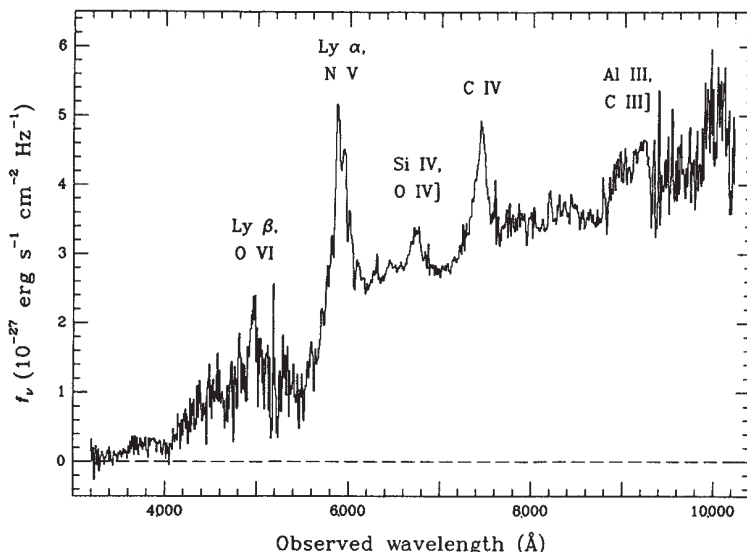


Fig. 1 Spectrum of 1208+1011, obtained with the Palomar Hale telescope on 13 January 1986 UT. A convenient continuum reference point¹³ is $m(6,960 \text{ Å}) \approx 17.8$, but the flux scale is only approximate (± 0.6 mag). Prominent features are marked. The blue spectrum ($\lambda \leq 5,800 \text{ Å}$) is affected by numerous Ly- α absorption lines. Incomplete removal of atmospheric absorption and emission lines produces narrow spikes in some places (such as the O $_2$ band at $\sim 6,860$ and $7,600 \text{ Å}$).

clouds in the range $2.5 \leq z \leq 3.8$ along the line of sight to 1208 + 1011. At wavelength longer than $\sim 6,000 \text{ \AA}$, the observed continuum can roughly be described by a power law, $f_\nu \propto \nu^{-1}$.

We thank J. Henning, C. Price, S. Staples and D. Tennant for assistance at Palomar Observatory. This research was supported by NSF grant AST84-16704 (W.L.W.S.) and by the Miller Institute for Basic Research in Science (A.V.F.).

Received 18 February; accepted 9 May 1986.

1. Hazard, C., McMahon, R. G. & Sargent, W. L. W. *Nature* **322**, 38-40 (1986).
2. Oke, J. B. & Gunn, J. E. *Publ. astr. Soc. Pacif.* **94**, 586-594 (1982).
3. Shectman, S. A. in *Annual Report of the Director, Mt Wilson and Las Campanas Observatories*, 676-677 (Carnegie Institution of Washington, 1982).
4. Filippenko, A. V. & Sargent, W. L. W. *Astrophys. J. Suppl. Ser.* **57**, 503-522 (1985).
5. Filippenko, A. V. & Greenstein, J. L. *Publ. astr. Soc. Pacif.* **96**, 530-536 (1984).
6. Oke, J. B. & Gunn, J. E. *Astrophys. J.* **266**, 713-717 (1983).
7. Lynds, C. R. *Astrophys. J.* **164**, L73-L78 (1971).
8. Young, P. J., Sargent, W. L. W., Boksenberg, A., Carswell, R. F. & Whelan, J. A. J. *Astrophys. J.* **229**, 891-908 (1979).
9. Sargent, W. L. W., Young, P. J., Boksenberg, A. & Tytler, D. *Astrophys. J. Suppl. Ser.* **42**, 41-81 (1980).
10. Young, P., Sargent, W. L. W. & Boksenberg, A. *Astrophys. J.* **252**, 10-31 (1982).
11. Peterson, B. A., Savage, A., Jauncev, D. L. & Wright, A. F. *Astrophys. J.* **260**, L27-L29 (1982).
12. Baldwin, J. A. in *Active Galactic Nuclei* (eds Hazard, C. & Mitton, S.) 51-78 (Cambridge University Press, 1979).
13. Baldwin, J. A. *Astrophys. J.* **214**, 679-684 (1977).
14. Phillips, M. M. & Hawley, S. A. *Publ. astr. Soc. Pacif.* **90**, 650-651 (1978).
15. Arp, H. *Astrophys. J.* **283**, 59-63 (1984).
16. Ulrich, M. H. & Perryman, M. A. C. *Mon. Not. R. astr. Soc.* **220**, 429 (1986).

The rotation period of Uranus

M. D. Desch, J. E. P. Connerney & M. L. Kaiser

Laboratory for Extraterrestrial Physics, Goddard Space Flight Center, Greenbelt, Maryland 20771, USA

The recent fly-by of Uranus by the Voyager 2 spacecraft provided a singular opportunity to measure one of the fundamental but poorly known physical properties of the planet, its intrinsic rotation period. Earth-based photometric and spectroscopic estimates of the 'atmospheric' period vary greatly, with values ranging from ~ 12 to 24 h (refs 1-4); estimates of the period based on the dynamical flattening of a rotating body range from ~ 15 to 17 h (refs 5, 6). Here we use the Voyager planetary radioastronomy⁷ and magnetometer⁸ observations at Uranus to derive a period of 17.24 ± 0.01 h. The greatly improved precision of this measurement provides useful constraints on models of the planet's internal structure.

The Voyager 2 radioastronomy data provide only an indirect determination of the rotation period, because we must make the commonly accepted assumption that the periodic nature of the emission is closely related to the intrinsic rotation of the planet. In order to minimize possible measurement bias due to unknown variations in beaming of the radiation with uranographic latitude, we restricted the analysis interval to post-encounter observations, when the spacecraft latitude changed by only a few degrees. This yielded a total data span of 42 planetary rotations, from 16:00 (spacecraft time) on 25 January to 20:00 on 24 February 1986, when we observed the last detectable emission from Uranus.

This is an extremely short interval compared with the nearly 600 rotations used to determine Saturn's period⁹. However, the discovery⁷ by Voyager of a highly repetitive emission feature has permitted a far more precise determination than would have been possible for Saturn using an equivalent span of data. The clock-like character of this feature is apparent in total energy plots (Fig. 1), in which the radio-wave energy is derived from an integration over the frequency band from 40 to 800 kHz. The highly periodic radiation drop-outs are evident as decreases of 2-3 orders of magnitude in the total radiated power.

To obtain an initial estimate of the rotation period, the data were Fourier-analysed, following correction of the observation

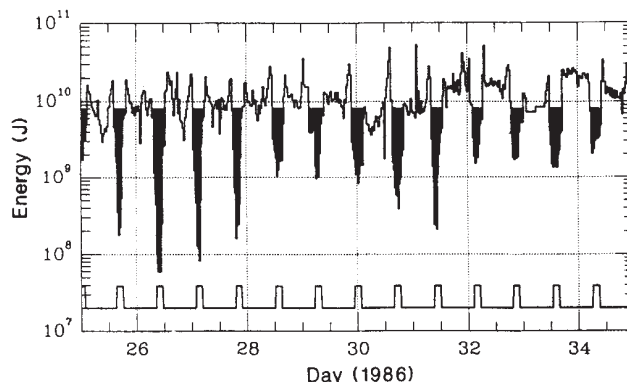


Fig. 1 Energy emitted by the Uranus radio source as a function of time for 10 days (14 rotations). For clarity, the emission drop-outs are shaded. An artificial 17.24-h clock is shown as a square wave pulse in the bottom of the figure. The highly periodic nature of the emission drop-outs is evident from a comparison of the timing of the event minima with the timing of the 17.24-h clock.

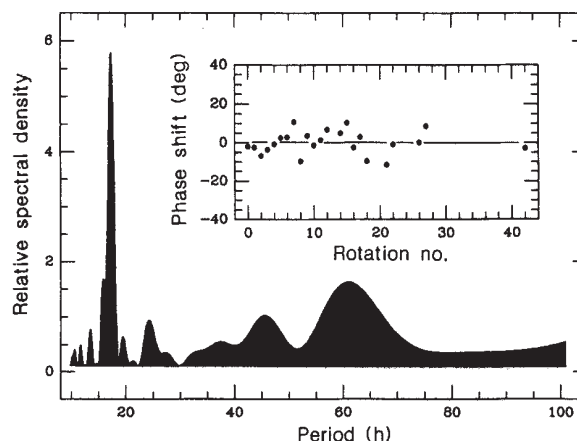


Fig. 2 Power spectral analysis of the Uranus radio-energy data of Fig. 1. The primary spectral peak is between 17.0 and 17.5 h. A linear least-squares fit to the midpoint times of the emission drop-outs measured over 42 Uranus rotations is shown in the inset. When the midpoints are plotted in an arbitrary longitude system that rotates with a period of 17.24 h, the result is zero average phase shift, representing no long-term drift of the events with time. The maximum deviation of individual points from the least-squares line is $\sim \pm 12^\circ$, corresponding to events leading or lagging a perfect clock by ± 35 min.

times for spacecraft motion and differential light travel time. Figure 2 shows the resulting power spectral density as a function of period. There is only one significant peak in the power spectrum, corresponding to the Uranus rotation period, at 17.25 ± 0.25 h.

This determination was improved by performing a linear least-squares fit to the midpoint times of the emission drop-outs. The procedure was very similar to that used to derive Saturn's rotation period⁹. By computing the phase of each midpoint in an arbitrary longitude system, the period and hence the phase shift of each point was adjusted until negligible long-term drift (zero least-squares slope) of the events with time was achieved (see Fig. 2 inset). This yielded a rotation period of 17.239 ± 0.006 h.

A second source of error in determining the period from the radioastronomy data arises from the fact that the midpoints were non-randomly distributed about the best-fit line (see Fig. 2 inset). This systematic variation introduced an additional uncer-

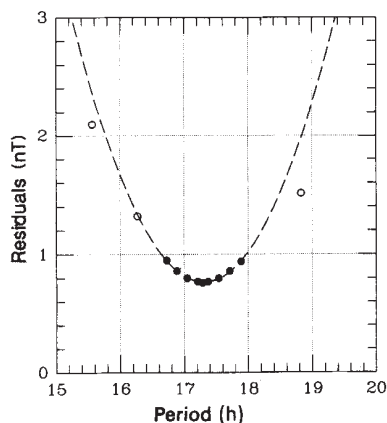


Fig. 3 Root-mean-square residual error between the best-fitting magnetic field model and the observations, as a function of assumed rotation period. The minimum r.m.s. error was obtained for a rotation period of 17.29 h. The parabolic least-squares line, shown by the dashed curve, is fit to the filled points.

tainty in the period measurement of ± 0.003 h. The final determination from the radioastronomy data is, therefore, 17.239 ± 0.009 h.

The magnetic field data provide a direct measure of the rotation period of the planet's interior, where the field is generated. The data used in this study spanned the interval from 15:00 to 21:00 (spacecraft time) on 24 January 1986. During this time Voyager 2 traversed the innermost magnetosphere, passing from a radial distance of $8 R_U$ (the planet radius $R_U = 25,600$ km) at 40° S latitude, through closest approach ($4.2 R_U$), to a radial distance of $8 R_U$ at 78° N latitude. The maximum observed magnetic field, measured just before closest approach, was 413 nT. Without accurate knowledge of the planet's rotation period, the location of the spacecraft in the reference frame rotating with Uranus is unknown, making global magnetic field determinations difficult. In order to determine the rotation period, spherical harmonic analyses were performed with the rotation period treated as a free parameter. For each assumed period, ranging from 15.5 to 18.8 h, the best-fitting I2E1 (dipole plus quadrupole internal field plus uniform external field) model was found.

Figure 3 shows the root-mean-square (r.m.s.) residual between the best-fitting models and the observations as a function of assumed period. The best fit was obtained for a rotation period of 17.29 h, for which the r.m.s. error was 0.76 nT. We believe that this residual error is largely due to that part of the external fields (for example, fields due to the magnetopause and magnetotail currents) not well approximated by a uniform field. Figure 3 also shows a parabolic least-squares fit to the points having r.m.s. error < 1 nT; the uncertainty in the parabola coefficients yields an uncertainty of ± 0.05 h in the rotation period.

The I2E1 model is the simplest spherical harmonic model for which an acceptable fit to the observations is obtained. Centred dipole models fit poorly (r.m.s. error > 7.8 nT) for all assumed rotation periods. An offset, tilted dipole representation of the field, presented in ref. 8, has an associated r.m.s. error of 2.4 nT. More complex (for example, octupole) models can be considered only in part using generalized inverse methods¹⁰, because of the limited observations. An examination of the impact of increasing model complexity on the deduced rotation rate suggests an uncertainty several times that obtained assuming an I2E1 model, or $\sim \pm 0.2$ h. The best direct estimate of the rotation period is then 17.29 ± 0.2 h.

The two period determinations made here are entirely independent and self-consistent. Combining the two periods yields a weighted mean value of $P = 17.24 \pm 0.01$ h. As no further

spacecraft observations of Uranus are planned at this time, we believe that this value is not likely to be improved upon in the foreseeable future.

The 17.24-h rotation period has important consequences for studies of atmospheric dynamics¹² and the internal structure and composition of Uranus. Inferences regarding the internal structure can be drawn from the relationship between the observed planetary oblateness, rotation period and gravitational moment (J_2). The latter two are now known with great precision, constraining plausible models of the interior. Several 2- and 3-layer models, incorporating a 'rock' core, 'ice' mantle and gaseous outer envelope, have been considered^{13,14} for Uranus. A model (U4, ref. 14) containing a rocky core of ~ 6.6 Earth masses ($M_\oplus$) and an outer envelope with $\sim 3.7 M_\oplus$ H_2 -H gas and $4.4 M_\oplus$ ice is most consistent with the rotation period reported here.

We thank our colleagues on the Voyager magnetometer investigation, led by principal investigator N. Ness; our colleagues on the Voyager planetary radioastronomy investigation, led by J. Warwick; and the data processing and display staff of the Laboratory for Extraterrestrial Physics, GSFC, under the direction of W. Mish, for support during the intensive Uranus encounter.

Received 23 April; accepted 16 May 1986.

1. Brown, R. A., Goody, R. *Astrophys. J.* **235**, 1066-1070 (1980).
2. Trauger, J. T., Roessler, F. L. & Munch, G. *Astrophys. J.* **219**, 1070-1083 (1978).
3. Trafton, L. *Icarus* **32**, 402-412 (1977).
4. Munch, G. & Hippelein, H. *Astr. Astrophys.* **81**, 189-197 (1979).
5. Elliot, J. L. *et al. Astr. J.* **86**, 444-455 (1981).
6. Goody, R. in *Uranus and the Outer Planets* (ed. Hunt, G.) (Cambridge University Press, 1982).
7. Warwick, J. W. *et al. Science* (in the press).
8. Ness, N. F. *et al. Science* (in the press).
9. Desch, M. D. & Kaiser, M. L. *Geophys. Res. Lett.* **8**, 253-256 (1981).
10. Connerney, J. E. P. *J. geophys. Res.* **86**, 7679-7693 (1981).
11. Stone, P. H. *Icarus* **24**, 292-298 (1975).
12. Smith, B. *et al. Science* (in the press).
13. Hubbard, W. B. & MacFarlane, J. J. *geophys. Res.* **85**, 225-234 (1980).
14. Podolak, M. & Reynolds, R. T. *Icarus* **46**, 40-50 (1981).

Detection of stratospheric HNO_3 and NO_2 response to short-term solar ultraviolet variability

G. M. Keating*, J. Nicholson III†, G. Brasseur‡, A. De Rudder‡, U. Schmailzl§ & M. Pitts||

* Atmospheric Sciences Division, NASA Langley Research Center, Hampton, Virginia 23665-5225, USA

† Vigyan Research Associates, Hampton, Virginia 23666, USA

‡ Institut d'Aéronomie Spatiale, 1180 Brussels, Belgium

§ Max Planck Institute for Chemistry, 6500 Mainz, FRG

|| SASC Technologies, Hampton, Virginia 23666, USA

Variations in the solar ultraviolet irradiance with a period equal to or approximately one-half of the rotation period of the Sun (27 days) are currently observed by satellite monitoring. These variations have typical peak-to-peak amplitudes of $\sim 3\%$ between 170 and 208 nm and $< 1.5\%$ at longer wavelengths. Detection of the response of stratospheric species to solar ultraviolet variability is crucial for understanding the photochemical behaviour of the middle atmosphere. Understanding the natural variations of the stratosphere is also a prerequisite for isolating possible anthropogenic effects; up to now most work has concerned stratospheric ozone¹⁻⁸. Recent analysis⁵ of measurements by the Nimbus 7 satellite's LIMS (limb infrared monitor of the stratosphere) and SBUV (solar backscatter ultraviolet) experiments has established very high correlations between ozone mixing ratios (detrended and corrected for temperature effects) and short-term variations in 205-nm solar radiation. A similar approach is used here to detect

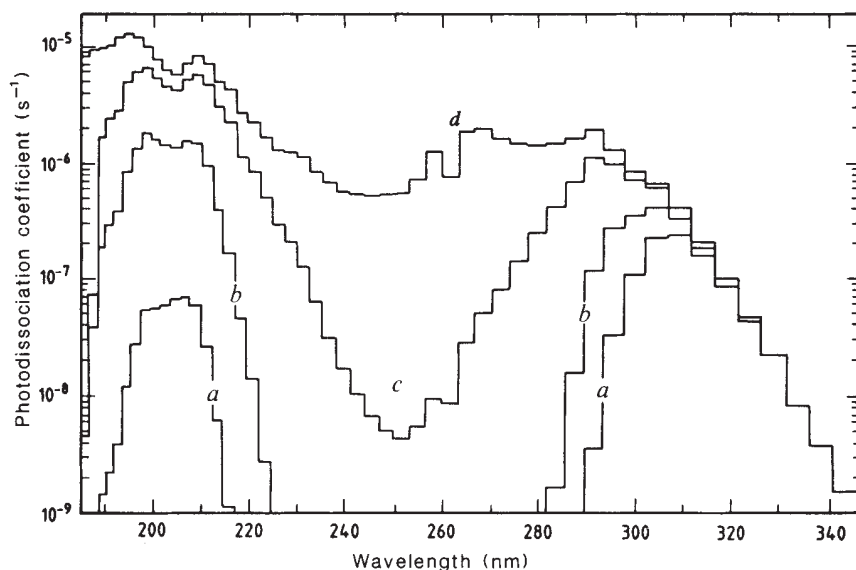


Fig. 1 Spectral distribution between 185 and 345 nm of the photodissociation coefficient of nitric acid for selected altitudes: *a*, 20 km; *b*, 30 km; *c*, 40 km; *d*, top of the atmosphere. The calculation assumes an overhead Sun (from ref. 16).

a relation between LIMS measurements of HNO_3 ^{9,10} and NO_2 ¹¹ and the SBUV measurements of short-term variations in 205-nm radiation¹². Observations show that the response of HNO_3 is much stronger than, but in the opposite sense to the ozone response, and that the NO_2 response is in the opposite sense to the HNO_3 response. This is the first detection of such a response for species other than ozone. Model calculations are in fair agreement with observed short-term variations and predict large variations in HNO_3 over the 11-yr solar cycle.

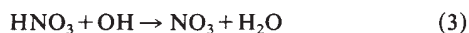
Nitrous oxide (N_2O) is the major source of odd nitrogen ($\text{N} + \text{NO} + \text{NO}_2 + \text{NO}_3 + 2\text{N}_2\text{O}_5 + \text{HNO}_3 + \text{HO}_2 + \text{HO}_2\text{NO}_2 + \text{ClONO}_2$) in the stratosphere. It is formed in the soil by bacterial action and transported into the middle atmosphere, where it is destroyed by photodissociation or by reaction with the electronically excited oxygen atom ($\text{O}(^1\text{D})$). Nitric oxide (NO), which is produced by the reaction of N_2O with $\text{O}(^1\text{D})$, is converted to nitrogen dioxide (NO_2) essentially through reaction with ozone, but also by reaction with ClO , HO_2 or CH_3O_2 . The NO_2 molecule reacts rapidly during the day to form NO either by reaction with O or by photolysis. During the night a significant amount of NO_2 reacts with O_3 to produce NO_3 , which, in turn, is converted into N_2O_5 by reaction with NO_2 . After sunrise the N_2O_5 concentration decreases continuously as solar ultraviolet radiation (primarily above 300 nm) photodissociates the molecule, replenishing NO_2 . The behaviour of NO_2 is also affected by chlorine compounds, as ClONO_2 is formed by the three-body reaction of ClO and NO_2 . The ClONO_2 is partly destroyed during the day by photodissociation or reaction with O or OH . Nitric acid is formed by the three-body reaction



The HNO_3 molecule is relatively stable in the lower stratosphere and is therefore subject to a variety of transport processes leading, for example, to large mixing ratios at high latitudes. The loss processes for HNO_3 are



and



The sensitivity of HNO_3 to change in the solar ultraviolet output is predicted by a one-dimensional time-dependent chemical-radiative model^{13,14}. The short-term spectral variability⁵ of the solar irradiance is applied in the model in accord with the solar data of Heath *et al.*¹². The average solar ultraviolet irradiances adopted in the model are based on ref. 15, and we assume a sinusoidal variation of the solar ultraviolet output.

At 10 mbar, the model predicts a reduction of 0.4% in HNO_3 for a 1% increase in solar radiation at 205 nm, with a response time of less than 1 day. This strong response to short-term solar ultraviolet variations can be understood by referring to Fig. 1 (see ref. 16), in which the photodissociation frequency of HNO_3 at different altitudes is shown as a function of wavelength for overhead Sun conditions. The decrease of photodissociation with decreasing altitude for wavelengths centred on 250 nm is due to increasing ozone absorption. The sharp decrease below 200 nm is due to absorption in the molecular oxygen Schuman-Runge bands. The net effect of solar absorption by stratospheric O_3 and O_2 results in an enhancement of the effects of solar variability near 200 nm on the photodissociation of HNO_3 .

The detection of the HNO_3 response to ultraviolet variations can be expected only in regions of the atmosphere where the chemical lifetime is short compared with the transport lifetime, and significantly shorter than the period of the solar oscillation. A calculation based on the model of Crutzen and Schmailzl¹⁷ shows that the chemical lifetime of HNO_3 increases with decreasing altitude and, at a given level, increases with latitude, particularly in winter. At 10 mbar near equinox, for example, the HNO_3 lifetime is enhanced by 40% going from 0 to 40° latitude, while at 20 mbar it is enhanced by 60% over the same latitudinal range. In winter, the variation of the lifetime at a given altitude is larger than in other seasons. At 10 mbar, for example, the lifetime increases by more than 300% between 0 and 40° latitude. Therefore, data from near the winter solstice and high-latitude data (latitudes > 40°) were removed from the LIMS data set. Moreover, the width of the latitudinal region about the Equator from which data were used was decreased at lower altitudes ($\pm 40^\circ$ at 10 mbar and $\pm 20^\circ$ at 20 mbar), to reduce dynamical effects. By this method, the data examined were limited to those from atmospheric regions where the photochemical lifetime of HNO_3 is calculated to be less than 2 days.

Finally, to remove any possible bias, the measurements were detrended relative to a running mean, averaged over a period slightly longer than the solar oscillation period characteristic of the interval investigated (similar to the approach used for the ozone data by Keating *et al.*⁵). No significant HNO_3 -temperature relation was predicted or observed at this level, so no correction was made. At 10 mbar, Fig. 2 shows the resulting relation between the ultraviolet ratio $(I_{205} - \bar{I}_{205})/\bar{I}_{205}$ and the dayside HNO_3 ratio $(\text{HNO}_3 - \bar{\text{HNO}}_3)/\bar{\text{HNO}}_3$, where I_{205} is the 5-day running mean of the 205-nm solar radiation, $\bar{I}_{205}$ is the 19-day running mean of I_{205} , HNO_3 is the 5-day running mean of the nitric acid volume mixing ratio averaged zonally and averaged between +40° and -40° latitude, and $\bar{\text{HNO}}_3$ is the

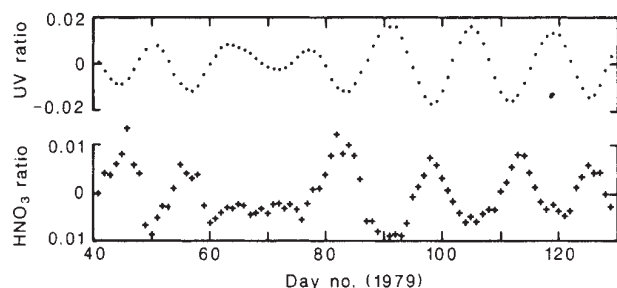


Fig. 2 Relation between daytime nitric acid ratio $[(\text{HNO}_3 - \text{HNO}_3)/\text{HNO}_3]$ at 10 mbar (+) and 205-nm solar irradiance (UV) ratio $[(I_{205} - \bar{I}_{205})/\bar{I}_{205}]$ (●) between days 40 and 130 of 1979. The correlation coefficient between the two parameters is $r = -0.78$.

19-day running mean of HNO_3 . Clearly, HNO_3 is negatively correlated with the approximate 13.5-day solar ultraviolet oscillation. The correlation coefficient between the two parameters shown in Fig. 2 is -0.78 . This correlation with the 205-nm solar irradiance is significantly higher than the correlation (-0.38) with the emission from the Sun at 10.7 cm (the classical 10.7-cm solar index). This is to be expected because, as indicated previously, the 205-nm radiation is very efficient in photodissociating HNO_3 throughout the stratosphere. The regression coefficient (sensitivity) for the 10-mbar daytime HNO_3 ratio against the ultraviolet ratio is -0.44 ± 0.07 (2σ uncertainty). In order to determine the lag time between the solar variability and the nitric acid response, the correlation coefficient between the two parameters was determined as a function of positive and negative lag (days). The resulting relationship is shown in Fig. 3: the maximum correlation occurs for a lag of 0 ± 1 days.

The 0.44% reduction in HNO_3 at 10 mbar can be compared with the variation of ozone at the same level. At low latitudes, the ozone response is calculated to be an increase of 0.17% for a 1% increase in 205-nm radiation, with a response time of 4 days (ref. 5). Observations indicate an increase in O_3 of $0.16 \pm 0.05\%$ for a 1% increase in 205-nm radiation, and a response time of 2 ± 1 days (ref. 5). Thus, the HNO_3 has an even stronger short-term response to solar ultraviolet variability than has ozone, and in the opposite sense.

It is difficult to study HNO_3 at higher altitudes because of the sharp decrease in concentration with altitude, but a similar study can be performed at 20 mbar over a smaller latitudinal interval ($\pm 20^\circ$) to minimize dynamical effects. Table 1 summarizes the results reported here and compares the model prediction (sensitivity (S) of HNO_3 , expressed in per cent change, to a 1% increase in the solar 205-nm irradiance and the calculated time lag of the response) to the corresponding values derived from the LIMS data. The statistical analysis has been performed using the combined day and night data, the day data, and the night data, respectively. As the model provides 24-h averaged values, only the results provided by the combined day and night data should be compared with the calculated sensitivities. As expected, the correlation coefficient (r) is higher for the daytime observations than for the night-time. A solar effect is, nevertheless, still visible in the night data, because the signature of the daytime photochemistry is essentially frozen during the night

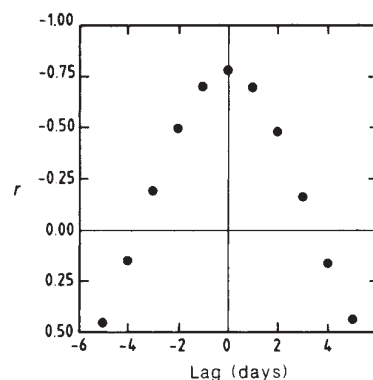


Fig. 3 Correlation coefficient between the nitric acid ratios and the 205-nm solar irradiance ratios in Fig. 2 for different time lags between these two quantities.

and is degraded principally by dynamics. The theoretical sensitivities calculated for the 13.5- and 27-day periods are in good agreement at 10 mbar with the value obtained from the data analysis, but are smaller than the observation at 20 mbar. One should, however, recall that the model provides globally-averaged sensitivities, whereas the data used for the 20 mbar analysis proceed from observations confined to the tropics, where sensitivities should be higher.

Because the photodissociation of HNO_3 leads to the formation of nitrogen oxides, there should also be a solar effect on the NO_2 mixing ratio. The LIMS NO_2 data, averaged in the same way as the HNO_3 data, show a rapid response time (less than 1 day) and positive correlation with the solar irradiance. At 10 mbar, an increase in combined day and night NO_2 of $0.22 \pm 0.12\%$ is found, for a 1% increase in the 205-nm solar irradiance. The model^{13,14} provides a corresponding increase of 0.15% in NO_2 for the 24-h average conditions. The observed NO_2 sensitivity on the dayside is larger ($+0.36 \pm 0.15\%$).

The variation in the nitric acid concentration in the stratosphere over the 11-yr solar cycle can be estimated by calculating the HNO_3 /ultraviolet sensitivity from a steady-state version of the model and by scaling the amplitude of the 205-nm solar forcing to the assumed 11-yr variation. If a $10 \pm 5\%$ increase is adopted for the 205-nm solar irradiance (based on a number of studies, including a regression analysis between detrended 205-nm and 10.7-cm solar irradiance over a period of 4 yr), the resulting long-term variation in the HNO_3 mixing ratio is estimated to be $-10 \pm 5\%$ at 10 mbar and $-5 \pm 2\%$ at 20 mbar. The sensitivity of the HNO_3 concentration to the solar ultraviolet irradiance is thus significantly increased for this longer-period oscillation. Such a large variation may be detectable in long-term measurements of HNO_3 .

The detection of a solar signal in the LIMS HNO_3 and NO_2 variations provides a new test of the validity of currently assumed photochemistry in the stratosphere. Indeed, the HNO_3 response to short-term solar variability depends on several simultaneous photochemical processes, including the sensitivity of the HNO_3 photodissociation rate to the solar radiation near 200 nm and the sensitivity of HNO_3 production to variations in NO_2 and OH. These latter parameters are modulated through O_3 and N_2O variations caused by solar ultraviolet changes. The

Table 1 Response of HNO_3 to 205-nm solar variability

Pressure (mbar)	Model prediction		Combined day and night		Day		Night		Δ lat.
	13.5 days	27 days	S	r	S	r	S	r	
10	-0.37	-0.42	-0.42 ± 0.07	-0.74	-0.44 ± 0.07	-0.78	-0.42 ± 0.12	-0.58	$\pm 40^\circ$
20	-0.13	-0.15	-0.29 ± 0.08	-0.57	-0.35 ± 0.09	-0.59	-0.25 ± 0.13	-0.34	$\pm 20^\circ$

Sensitivity, $S = \frac{(\text{HNO}_3 - \bar{\text{HNO}_3})/\text{HNO}_3}{(I_{205} - \bar{I}_{205})/\bar{I}_{205}}$; correlation coefficient, r . Time lag at both altitudes: 0.5 ± 0.5 days (model) and 0 ± 1.0 days (data analysis).

observed amplitude and phase of the responses at 10 mbar of HNO_3 and NO_2 to solar ultraviolet variability are in fair accord with predicted responses based on the currently accepted chemical scheme¹⁸ used in the model. The results also imply large variations of HNO_3 over the 11-yr solar cycle, which should significantly affect other chemical species, including ozone. Indeed, increased solar activity can result in increased conversion of HNO_3 into active NO_x , which in turn can affect the ozone balance.

We thank J. M. Russell III for providing data on HNO_3 , NO_2 and temperature from the Nimbus 7 LIMS experiment in November 1983, D. F. Heath for providing 205-nm solar spectral irradiance values from the Nimbus 7 SBUV experiment in August 1984, and J. E. Nealy for photochemical calculations. Some of this work was accomplished under NASA contract NAS1-15785 and CMA contract 83-468.

Received 18 February; accepted 11 March 1986.

1. Keating, G. M. *Sol. Phys.* **74**, 321-347 (1981).
2. Chandra, S. J. *geophys. Res.* **89**, 1373-1379 (1984).
3. Hood, L. L. *J. geophys. Res.* **89**, 9557-9568 (1984).
4. Gille, J. C., Smythe, C. M. & Heath, D. F. *Science* **225**, 315-317 (1984).
5. Keating, G. M., Brasseur, G. P., Nicholson III, J. Y. & De Rudder, A. *Geophys. Res. Lett.* **12**, 449-452 (1985).
6. Heath, D. F. & Schlesinger, B. M. in *Atmospheric Ozone* (eds Zerefos, C. S. & Ghazi, A.) (Reidel, Dordrecht, 1985).
7. Hood, L. L. *J. geophys. Res.* **91**, 5264-5276 (1986).
8. Chandra, S. J. *geophys. Res.* **91**, 2719-2734 (1986).
9. Gille, J. C. *et al. J. geophys. Res.* **89**, 5179-5190 (1984).
10. Russell III, J. M. *Adv. Space Res.* **4**, (4), 107-116 (1984).
11. Russell III, J. M. *et al. J. geophys. Res.* **89**, 5099-5107 (1984).
12. Heath, D. F., Donnelly, F. F. & Merrill, R. G. *NOAA tech. Rep. No. ERL 424-ARL7* (1983).
13. Brasseur, G., De Rudder, A., Keating, G. M. & Pitts, M. J. *geophys. Res.* (submitted).
14. Brasseur, G., De Rudder, A. & Roucou, A. in *Proc. int. Conf. Environmental Pollution*, 839-910 (University of Thessalonika, 1982).
15. Brasseur, G. & Simon, P. C. *J. geophys. Res.* **86**, 7343-7362 (1981).
16. Biau, F. J. *Photochem.* **2**, 139-149 (1973).
17. Crutzen, P. & Schmalz, U. *Planet. Space Sci.* **31**, 1009-1032 (1983).
18. Demore, W. B. *et al. Jet Propulsion Lab. Publ. No. 85-37* (1985).

Novel experimental investigation of spontaneous ignition of gaseous hydrocarbons

J. F. Griffiths & W. Nimmo

Department of Physical Chemistry, The University, Leeds LS2 9JT, UK

Studies of spontaneous ignition of gases under rapid compression¹⁻³ are directed mainly towards understanding 'knock' in spark-ignition engines and the duration of ignition delays in diesels. The relevance to engine combustion lies in the ability to investigate the effects of fuel structure on the development of spontaneous ignition (arising either in different isomers of a hydrocarbon or due to reaction of mixtures of different hydrocarbons), while avoiding complicating practical features such as droplet evaporation, induction and exhaust of the charge, reciprocating piston motion and cycle-to-cycle variations. We report here a novel development, in which a variable-speed impeller in the combustion chamber of a rapid compression apparatus is used, first, to bring about spatial uniformity of temperature and concentration throughout the ignition delay period and, second, to enhance the rate of heat dissipation by successive increases in the speed of rotation. We show that, at given conditions for rapid compression, spontaneous ignition ceases to be possible beyond a limiting rotor speed.

The single-shot rapid compression machine⁴ is operated at a compression ratio ($V_{\text{initial}}/V_{\text{final}}$) of 10.6:1 and at a combustion chamber temperature (T_a) of 293 K. During operation the piston

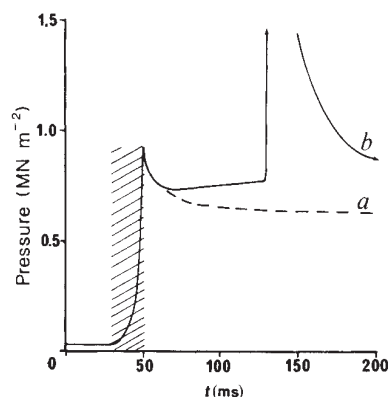


Fig. 1 Pressure-time records for non-reactive (a) and reactive (b) compositions compressed rapidly under identical conditions in the absence of mechanical stirring. Ignition occurs after a delay of 80 ms from the time when rapid piston motion ceases. The shaded area represents the interval during which rapid piston motion takes place. The zero of time is selected by the microcomputer control system.

travels 23.2 cm from its initial position in ~22 ms and on completion of its stroke has compressed the pre-loaded (33.0 kN m^{-2}), pre-mixed gaseous reactants into a squat cylindrical chamber of volume 31.9 cm^3 and depth 1.92 cm. End-of-compression pressures are $\sim 0.9 \text{ MN m}^{-2}$ ($\sim 9 \text{ atm}$) yielding, at the moment the piston stops, gas temperatures (T_c) of 775 K in the present study. A virtually ideal adiabatic compression is achieved, and so this temperature is determined by the relationship $T_c/T_a = (V_i/V_f)^{\gamma-1} = (P_i/P_f)^{1/\gamma}$, where γ is the ratio of the principal specific heats of the reactant mixture. Our choice of initial composition, $0.03 \text{ nC}_4\text{H}_{10} + 0.20 \text{ O}_2 + 0.77 \text{ Ar}$, is governed by the desire to achieve the highest possible initial temperature, using the present mechanical system, in reactants that mimic 'typical' hydrocarbon-air mixtures in engines. Thus a stoichiometric proportion of fuel to oxygen is used, but the nitrogen is substituted by argon so as to augment γ as far as possible. The highest value for γ is that of the monatomic gases ($\gamma = 5/3$), but only in this limit is γ independent of temperature. In practice, for mixtures of gases, γ is derived at the compression temperature by an iterative procedure using a second-order polynomial representing the molar heat capacity as a function of temperature for each component of the system.

Butane, in its isomeric *n*- and iso-forms, is the simplest hydrocarbon with which the effect of structural change on the spontaneous ignition behaviour of hydrocarbons can be investigated. Its gaseous state is preserved throughout the course of compression. This report concerns only *n*-butane.

Pressure changes that take place during compression, and afterwards in the closed, constant-volume chamber, are measured by a pressure transducer and recorded digitally. Because there is no significant loss of material by leakage and because the reactant is in sufficiently great dilution that only a very small change in the number of moles of material occurs during reaction, any pressure change measured during the post-compression period reflects a temperature change. Thus, in non-reactive compositions (Fig. 1, curve a) the gas temperature (and thus pressure) falls from its maximum at the end of piston travel due to heat loss to the walls. The pressure record can be used to calculate a mean gas temperature when heat transfer occurs only by residual gas motion due to the preceding piston motion. When vigorous mechanical stirring is introduced, the measured pressure is directly proportional to the spatially uniform gas temperature. Quantitative interpretations of heat release and dissipation rates may then follow³.

Curve b in Fig. 1 is the pressure record obtained, in the absence of stirring, in a reactive composition of the same heat

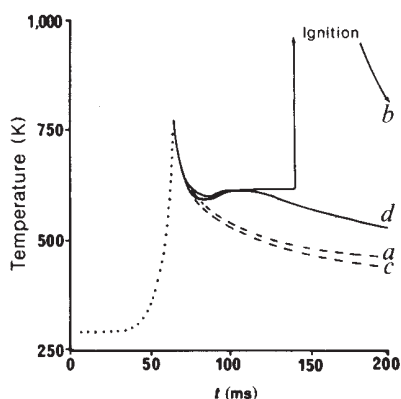


Fig. 2 Temperature-time records for non-reactive and reactive compositions following rapid compression with different rotor speeds. The dotted line represents the temperature change during the mechanical compression stroke. *a*, Non-reactive, stirring rate 850 ± 50 r.p.m.; *b*, reactive, stirring rate 850 ± 50 r.p.m.; *c*, non-reactive, stirring rate $1,100 \pm 50$ r.p.m.; *d*, reactive, stirring rate $1,100 \pm 50$ r.p.m.

capacity as that associated with curve *a*. When piston motion ceases, the two curves at first show an identical cooling rate: the initial rate exceeds 3×10^4 K s⁻¹. The separation of curve *b* from curve *a* at 10 ms after compression shows that heat is being released from the exothermic oxidation of butane at a sufficient rate to augment the gas temperature. The mean temperature corresponds to 610 K at this time. During the next 70 ms the temperature increases monotonically (by ~ 50 K): thermal feedback is taking place. There is a critical transition to ignition 80 ms after rapid piston motion ceases, due not only to the enhanced reaction rate caused by the rising temperature but also to chain branching by means of organic peroxides formed as molecular intermediates during the ignition delay period⁵. Photographic records of the development of ignition in the chamber have been published elsewhere⁶.

Figure 2 shows temperature-time records derived from pressure histories obtained under the effect of mechanical stirring. The dashed curves (Fig. 2*a, c*) are from the cooling of a non-reactive composition when the rotor speed is raised from 850 (± 50) to 1,100 (± 50) revolutions per minute. An enhanced cooling rate is discerned after 10 ms into the post-compression period, once the initially dominant gas movement caused by rapid piston motion has subsided sufficiently. The corresponding records from reactive compositions at the respective rotor speeds are shown as solid curves (Fig. 2*b, d*). There is a failure of spontaneous ignition at the higher rotor speed (curve *d*) even though exothermic oxidation has taken place. Ignition has failed in this case because the heat release rate, which rises to a maximum and then begins to wane as the primary reactant concentration falls, is unable to maintain the gas temperature. The heat release rate diminishes to zero at 80 ms after compression (W.N., unpublished data). At marginally supercritical conditions (curve *b*), the heat release also falls virtually to zero just before hot ignition takes place, but as a consequence of the lower heat dissipation rate, the reactant temperature remains higher and chain branching becomes effective, giving rise to chain-thermal criticality⁷.

We thank SERC, EEC and UKAEA for financial support.

Received 20 February; accepted 18 April 1986.

1. Martinengo, A. & Homann, K. H. *Oxidation and Combustion Reviews* Vol. 2 (ed. Tipper, C. F. H.) 207-256 (Elsevier, Amsterdam, 1967).
2. Fish, A., Haskell, W. W. & Reed, I. A. *Proc. R. Soc. A* **313**, 261-297 (1969).
3. Griffiths, J. F. & Hasko, S. M. *Proc. R. Soc. A* **393**, 371 (1984).
4. Beeley, P., Griffiths, J. F. & Gray, P. *Combust. Flame* **39**, 255-268; 269-282 (1980).
5. Griffiths, J. F. *Adv. chem. Phys.* **64**, 203-303 (1986).
6. Griffiths, J. F. & Nimmo, W. *Combust. Flame* **60**, 215-218 (1985).
7. Gray, B. F. *Trans. Faraday Soc.* **65**, 1623-1613 (1969).

Climatic change and Aboriginal burning in north-east Australia during the last two glacial/interglacial cycles

A. P. Kershaw

Department of Geography, Monash University, Clayton, Victoria, Australia 3168

Long palynological records from continental deposits may be divided into two categories: detailed sequences seldom extending back much further than the most recent interglacial¹⁻³, and more generalized or discontinuous sequences which cover all or a substantial part of the Quaternary⁴⁻⁶. I present here a record which is unusual in that it provides, in some detail, changes through a period considered to embrace the last two glacial/interglacial cycles. It provides the opportunity to compare the results of climatically-induced changes at corresponding stages within the two cycles and also to assess the impact of Aboriginal people on the vegetation. People have been present in Australia for the past 40,000 years⁷ and possibly as long ago as the last interglacial period⁸, but are unlikely to have been present before this.

The record presented here is an extension of that described previously from Lynch's Crater on the Atherton Tableland⁹⁻¹². Figure 1 shows selected results, including percentages of major gymnosperm and angiosperm big-tree taxa derived from rain-forest, and the most common canopy tree taxa from open sclerophyll vegetation, together with a summary of the relative proportions of all pollen from these three groups. Six major zones are recognized, of which the top four have been described previously¹¹. Zones G, E and A are dominated by pollen from rain-forest angiosperms, zones F and D contain high percentages of the rain-forest gymnosperms *Araucaria* and *Podocarpus* together with a substantial sclerophyll component, and zone B is dominated by the sclerophyll taxa. Zone C is transitional between rain-forest and sclerophyll-dominated periods. A timescale has been determined for the past 40,000 yr from radiocarbon dates and estimated beyond this time by applying known rates of accumulation of similar sediments from sites within the region to sections of the core, taking into account the degree of sediment compaction¹³.

A large degree of vegetation variation can be explained as a response to climatic change, particularly precipitation. Very high values for rain-forest angiosperms in zones G, E and A indicate domination by complex rain-forest existing in high rainfall conditions. Slightly higher gymnosperm values and higher percentages of *Elaeocarpus* relative to Cunoniaceae in subzone E1 than in zone A and subzones G2 and E3 suggest that precipitation did not reach equivalent levels, while generally higher sclerophyll and rain-forest gymnosperm percentages in subzones G1 and E2 indicate that effective precipitation was lower still.

Sharp reductions in precipitation are suggested at the beginning of zones F and D by substantial rises in sclerophyll taxa, particularly *Casuarina*, and the gymnosperm *Araucaria*. Subzone D2 can be separated on the basis of a higher rain-forest angiosperm component and a peak in *Elaeocarpus*, which signify slightly wetter conditions. The increases in sclerophyll taxa through zone C and their predominance in zone B are difficult to interpret in climatic terms, but precipitation was certainly no higher than in the period represented by zone D.

The best-estimate precipitation curve in Fig. 1 is based on information derived from modern pollen samples¹⁰ and on the known ecology of individual plant taxa and vegetation types. A more cautious precipitation estimate is shown as a range calculated mainly from bioclimatic analyses¹⁴ undertaken on an extensive Australian rain-forest data set¹⁵. The range is one standard deviation from the mean annual precipitation calculated for the most likely predominant structural vegetation type¹⁶

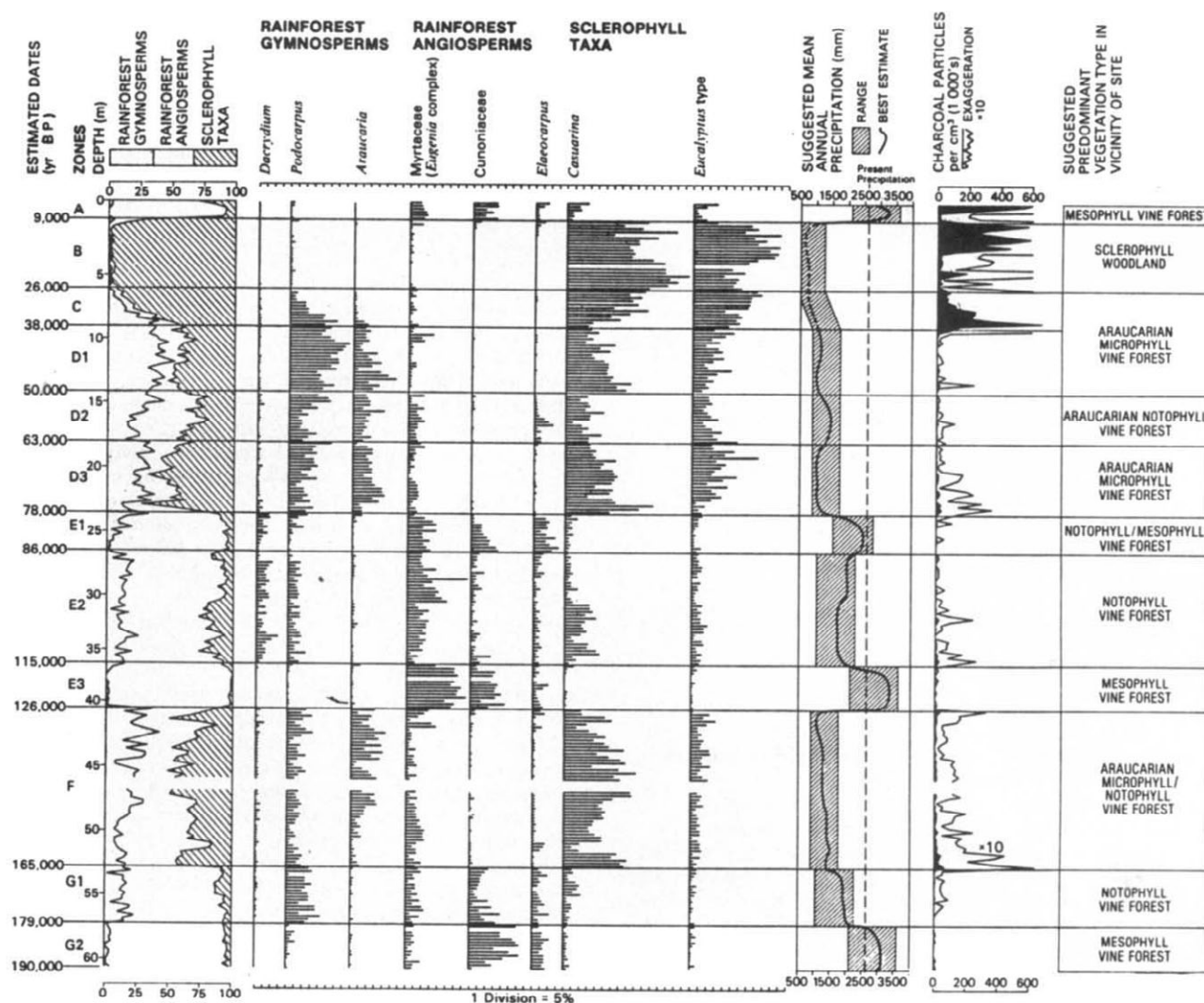


Fig. 1 Selected attributes of pollen diagram from Lynch's Crater. The frequencies of pollen of all taxa are shown as percentages of the dry-land-plant pollen total.

surrounding the pollen site during the period represented by each zone.

A comparison of the precipitation curve with that from the deep-sea core V28-238¹⁷ (Fig. 2) reinforces the suggestion from analyses of the more recent part of the record that precipitation changes within this area have been largely controlled by global changes in sea level and/or sea surface temperatures¹⁸.

Despite the strong climatic control of vegetation variation, there are some features of the record which cannot be explained by climate alone. One example is the inconsistent performance of *Eugenia* complex, particularly within the high-precipitation 'interglacial' periods represented by zones G2, E3 and A. This is at least partly related to the inclusion in the taxon of several different species with various ecological requirements and differential pollen production and dispersal capabilities. This is a common problem of palynology in areas dominated by highly diverse communities. A second feature is the attainment of somewhat different percentages by taxa and taxon groups in corresponding phases of the two identified cycles. This applies most obviously to the attainment of higher levels of gymnosperms within the penultimate glacial. The preferred explanation here relates to the general infilling of the lake basin and its effects on pollen influx. With the extension of swamp over the

lake basin, it is probable that the inwash pollen component, likely to have been dominated by lake-edge angiosperms with poor aerial pollen dispersal, would have been reduced. This would have effectively increased the proportions of wind-dispersed, more regional pollen such as that from the gymnosperms.

The most important difference between the two cycles is the almost total replacement of rain-forest by sclerophyll vegetation towards the end of the most recent dry period. It was originally proposed that a change in climate could not alone have brought about this change because of the potential of 'drier' rain-forest to survive mean annual precipitation levels as low as 600 mm on basaltic soils such as those surrounding the site, and because of the extinction of *Dacrydium* and perhaps other taxa which must have been equipped to survive the vicissitudes of Quaternary climates^{9,10}. Instead, an increase in burning, most probably due to the activities of Aboriginal people, was proposed as the most likely cause. Subsequently, charcoal analyses on the core sediment indicating a sharp increase in burning about 38,000 yr ago (see Fig. 1) gave some support to this interpretation⁸.

The evidence presented here, that sclerophyll vegetation did not totally replace rain-forest during the penultimate 'dry glacial' period in contrast to the last 'glacial', is strong support for a

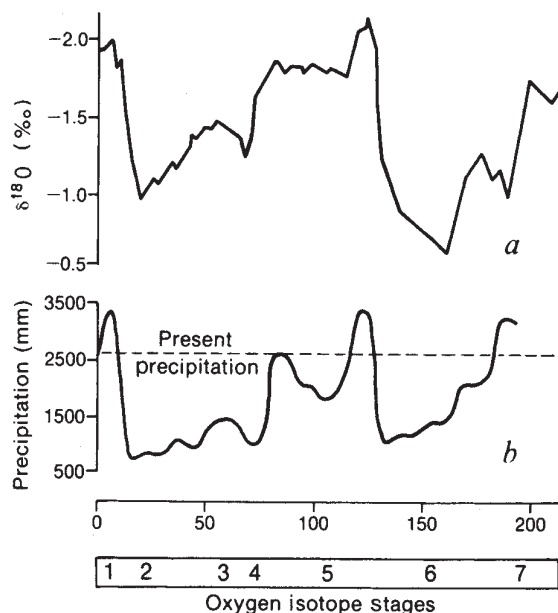


Fig. 2 Comparison of the precipitation curve derived from Lynch's Crater (b) with the oxygen isotope curve from deep-sea core V28-238 (a). Oxygen isotope stages from ref. 17.

recent increase in burning. Furthermore, the nature of the change from araucarian vine forest to sclerophyll vegetation can be shown to be different from other major changes in the record for which a climatic cause is invoked. Most changes from one zone to another are abrupt, and span perhaps only a few hundred years. Small charcoal peaks accompanying many of these changes suggest a role for fire in facilitating the climatically controlled transition from one vegetation state to another, as has been proposed for vegetation changes elsewhere¹⁹⁻²¹. By contrast, the change to sclerophyll vegetation is gradual, taking as long as the period represented by zone C, perhaps ~12,000 yr. It is best explained as a gradual replacement of fire-sensitive rain-forest by fire-promoting sclerophylls under a regime of frequent burning of the sclerophyll vegetation.

I thank C. Shaw, L. Head, S. Matiske, R. Phelps, I. Sluiter, G. Werren and Professor B. Thom for help with drilling, and the Australian Research Grants Scheme for finance.

Received 3 February; accepted 4 April 1986.

- Adam, D. P. & West, G. J. *Science* **219**, 168-170 (1983).
- Hauser, C. J. *Quat. Res.* **16**, 293-321 (1981).
- Woillard, G. M. *Quat. Res.* **9**, 1-21 (1978).
- Hooghiemstra, H. *Vegetation and Climatic History of the High Plain of Bogota, Colombia* (Cramer, Lichtenstein, 1984).
- Singh, G., Opdyke, N. D. & Bowler, J. M. *J. geol. Soc. Aust.* **28**, 435-452 (1981).
- Fuji, N. & Horie, S. *Proc. Jap. Acad.* **4**, 500-504 (1972).
- Pearce, R. H. & Barbetti, M. *Archaeol. phys. Anthropol. Oceania* **16**, 173-178 (1981).
- Singh, G., Kershaw, A. P. & Clark, R. in *Fire and Australian Biota* (eds Gill, A. M., Groves, R. A. & Noble, I. R.) 23-54 (Australian Academy of Sciences, Canberra, 1981).
- Kershaw, A. P. *Nature* **251**, 222-223 (1974).
- Kershaw, A. P. *New Phytol.* **77**, 469-498 (1976).
- Kershaw, A. P. *Nature* **272**, 159-161 (1978).
- Kershaw, A. P. *New Phytol.* **94**, 669-682 (1984).
- Kershaw, A. P. *Proc. Fourth int. Palynol. Conf.*, Lucknow 3, 28-35 (Birbal Sahni Institute of Palaeobotany, Lucknow, 1980).
- Nix, H. A. in *National Rainforest Study Report* (eds Werren, G. L. & Kershaw, A. P.) 421-425 (Geography Dept. Monash University, Clayton, 1984).
- Webb, L. J., Tracey, J. G. & Williams, W. T. *Aust. J. Ecol.* **9**, 169-198 (1984).
- Webb, L. J. *Aust. Plants* **9**, 349-363 (1978).
- Shackleton, J. J. & Opdyke, N. D. *Quat. Res.* **3**, 39-55 (1973).
- Kershaw, A. P. in *Quaternary Studies* (eds Suggate, R. P. & Creswell, M. M.) 181-187 (Royal Society of New Zealand, Wellington, 1975).
- Green, D. G. *J. Biogeog.* **9**, 29-40 (1982).
- Hookey, A. D., Southern W. & Kershaw, A. P. *J. Biogeog.* **7**, 349-362 (1980).
- Walker, D. in *The Plant Community as a Working Mechanism* (ed. Newman, E. I.) 27-43 (Blackwell, Oxford, 1982).

Application of laser holographic techniques to investigate crustal deformations

Shuzo Takemoto

Disaster Prevention Research Institute, Kyoto University, Kyoto, 611 Japan

Laser holographic techniques can be used to detect small deformations of three-dimensional objects, of the order of the wavelength of the laser light. We report the use of laser holography to measure the crustal deformation in a deep tunnel at Amagase, Kyoto, Japan. A recording system consisting of a helium-neon gas laser and associated optical elements was installed in the observation tunnel in 1984. Using this system, holograms of a section of the tunnel wall ~2 m in diameter are directly recorded on photographic plates. When the reconstructed image of a hologram is superposed on the current ('real-time') image of the tunnel wall, many interference fringes can be seen through the developed photographic plate. The fringe displacement, formed by the deformation of the tunnel, is continuously monitored with a video camera and a video-cassette recorder. The measured fringe displacement in the interference pattern is consistent with strain changes obtained from extensometers which have been installed in the same tunnel.

Observations of crustal deformation have been commonly made with extensometers by continuously measuring the relative displacement between two piers fixed into the bedrock. The

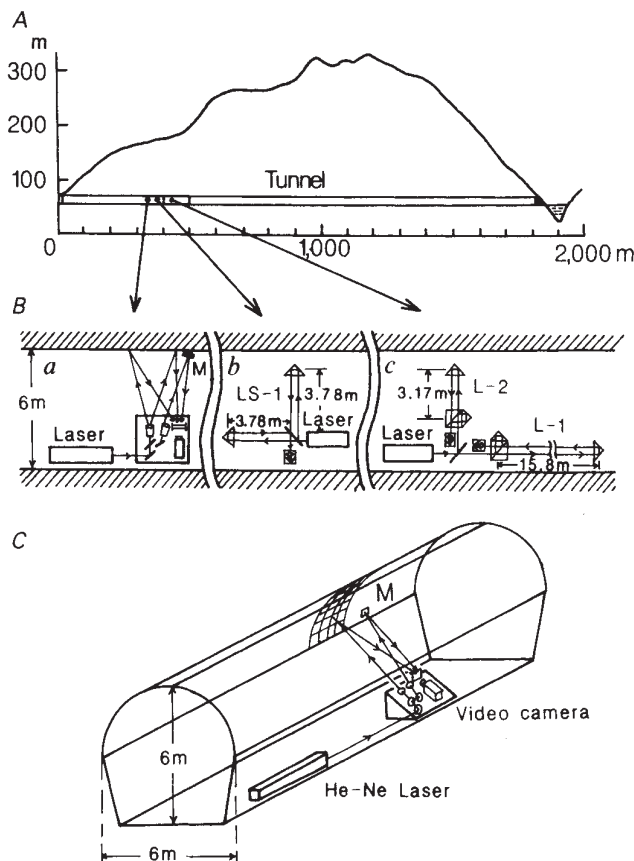


Fig. 1 A, Topographic profile of the Amagase tunnel. B, Arrangement of the laser holography system and laser extensometers installed in the tunnel. C, Isometric view of the laser holography system, comprising a He-Ne laser, optical mirrors and lenses, a photographic plate and a video camera.

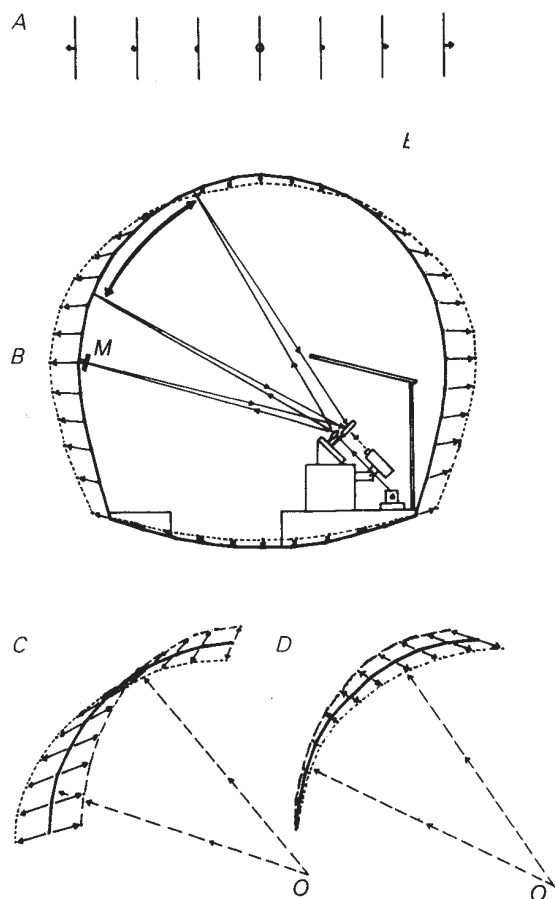


Fig. 2 Tunnel deformation calculated with the two-dimensional finite element method. *A*, Uniform strains in the absence of the tunnel due to a uniaxial tensile stress. *B*, Displacements around the horseshoe-shaped tunnel caused by the tensile stress which would produce uniform strains in the absence of the tunnel. *C*, Displacements of the tunnel wall relative to point *O*, where optical elements, except the reference mirror (*M*), are fixed. Dotted line, displacements due to tensile stress; dashed line, displacement due to compressive stress. *D*, Displacements of the tunnel wall relative to point *M*. Dotted and dashed lines as in *C*.

early developments of these instruments were due to Benioff¹, Ozawa² and Takada³. Their extensometers, using length standards of fused quartz tubes or super-invar bars, have been installed in many observatories for investigating secular strains, tidal strains and low-frequency seismic waves. Extensometers of this type, however, suffer from the disadvantages inherent in length standards of solid material, that is, deformation of the length standards themselves and frictional forces between length standards and their supports. In contrast, laser extensometers are free from these disadvantages because they are capable of detecting linear strains without using length standards of solid material. Various instruments have been developed by a number of groups in the past two decades⁴⁻¹¹. In addition, unlike conventional extensometric methods, which can detect only one-dimensional linear strains, the holographic method can detect two- or three-dimensional strain patterns.

Laser holography has been used to measure tunnel deformation at the Amagase Crustal Movement Observatory since 1984. A laser holographic recording system was installed 320 m from the entrance of a disused tunnel of length 1,830 m. The section of the tunnel has a horseshoe shape, with a diameter of ~6 m. The observation site is ~130 m below the surface. The annual variation of air temperature at the observation site is ~0.2 °C

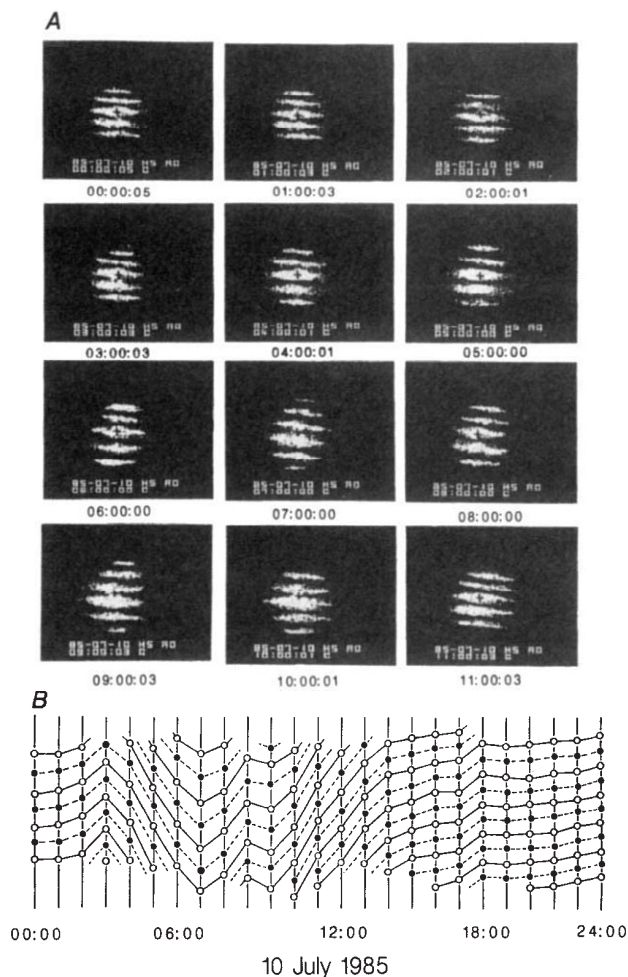


Fig. 3 *A*, Examples from 10 July 1985 of interference fringes superposed on the holographic image of the tunnel wall, on which the standard mark '+', has been drawn with black paint. *B*, Hourly plots of fringe displacements measured along the vertical line (~1 m) passing through the standard mark '+'. Open and filled circles represent bright and dark fringes, respectively. The plots are inverted (upside-down), for convenience of comparison with Fig. 4.

and its daily variation is <0.05 °C. The amplitude of microtremor measured in the tunnel does not exceed 20 nm s⁻¹ at ~0.1–10 Hz. These quiet circumstances allow us to obtain clear holograms of objects having a dimension of ~1 m. Figure 1 shows the arrangement of the holographic recording system together with three laser extensometers (L-1, L-2, LS-1) which have been installed in the same tunnel.

The holographic recording system consists of a helium-neon gas laser with an emission power of 50 mW, optical mirrors and lenses, a video camera and a video-cassette recorder. Optical elements, with the exception of a reference mirror, are fixed with magnetic holders to a massive steel plate (60 × 90 cm) which is fastened to a concrete base on the floor. The reference mirror ('M' in Fig. 1c) is attached to the tunnel wall with anchor bolts. The laser source is set on another concrete base and covered by polystyrene boards. A coherent light beam emitted from the laser source is split into two beams at the surface of the beam splitter, which deflects ~5% of the light beam and transmits the remaining portion. The former is used as a reference wave and the latter illuminates an area of the tunnel wall ~2 m in diameter. Two waves, reflected from the reference mirror and the tunnel wall, respectively, are superposed on a photographic plate (Agfa 10E75) measuring 6.5 × 9 cm. After an exposure of

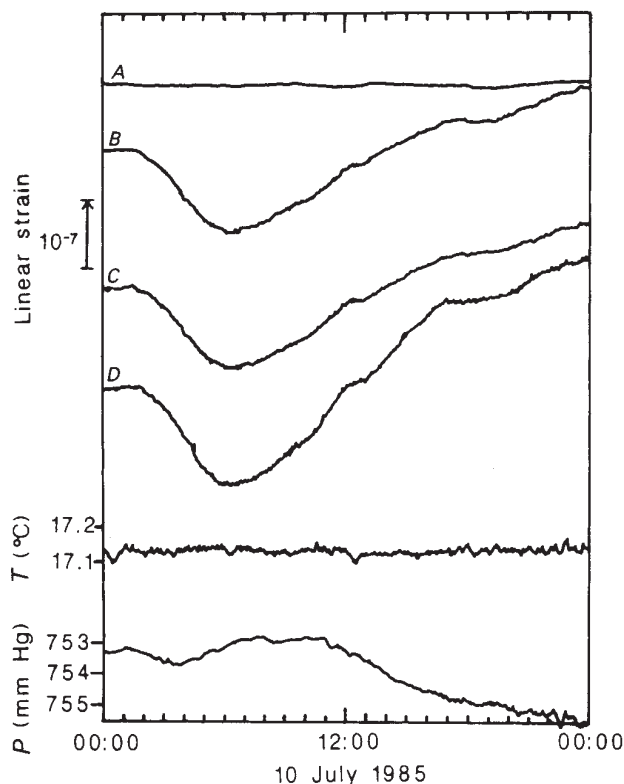


Fig. 4 Strain changes observed with laser extensometers installed in the Amagase tunnel [A, (L-1); B, (L-2); C, ((L-2)-(L-1)); D, (LS-1)], together with thermometric and barometric records on 10 July 1985.

~1-3 min, the photographic plate is developed and fixed. The developed plate is used as the original hologram, and is carefully reset in the position at which the hologram was taken. By again illuminating the plate with the same reference wave, the holographic image of the tunnel wall is reconstructed. In this procedure, the plate is slightly inclined from the original position. A number of dark and bright lines ('fringes') over the object image can then be seen on the hologram: these are caused by the interference of two waves; one reconstructed from the hologram and the other scattered directly by the tunnel wall. The movements of these interference fringes enable us to estimate the tunnel deformation.

We now calculate strain changes in a long tunnel which is assumed to be driven in an infinite, homogeneous, isotropic, elastic medium. Because the change in strain components along the tunnel axis can be assumed to be negligibly small compared with that across the tunnel, calculations are made with the two-dimensional finite element technique under the plane-strain condition. Results of the calculations are illustrated in Fig. 2B-D. Figure 2B shows displacements around the tunnel due to the uniaxial tensile stress which would produce uniform strains (Fig. 2A) in the absence of the tunnel; Fig. 2C and D show displacements of the tunnel wall relative to the points 'O' and 'M', respectively. Fringe displacements in the interference pattern at the Amagase tunnel are predicted from these figures. If the developed photographic plate is adjusted to produce interference fringes parallel to the direction of tunnel axis, upward and downward displacements of fringes and changes in their width should be observed: the former correspond to the mean displacement of the tunnel wall relative to point O and the latter correspond to the tilt of the tunnel wall relative to point M.

Figure 3 shows examples of fringe patterns recorded with the video-cassette recorder, together with hourly plots of their position on 10 July 1985. The drift of the fringe indicates an initial

upward trend, reversing at 03:00 and again at 07:00. The fringe width is narrowest at 03:00 and widest at 07:00.

The holographic results may be compared with the strain changes measured by laser extensometers.

Figure 4A and B are the linear strains measured by laser extensometers⁹ L-1 and L-2, with a frequency-stabilized laser source and mutually perpendicular evacuated light paths. As shown in Fig. 1B, L-1 is oriented along the tunnel and L-2 across the tunnel. The curve in Fig. 4C shows the difference between these two components of linear strain ((L-2)-(L-1)). The curve in Fig. 4D shows the strain measured by a new type of laser extensometer¹⁰ which has a simple, unstabilized laser source and two light paths of equal length, and detects the difference between two axial linear strains. (LS-1) is thus equivalent to (L-2)-(L-1). The strain change along the tunnel (L-1) is smaller by an order of magnitude than that across the tunnel, (L-2): this is explained by the cavity effect¹². Therefore, (L-2)-(L-1) and (LS-1) are nearly equal to (L-2), which has a maximum at 02:00 and a minimum at 07:00 on 10 July.

The trends of the strain changes observed with extensometers are consistent with those of the fringe displacements in the interference pattern shown in Fig. 3B. Thus, the fringe displacements observed with the laser holographic method are indicative of the tunnel deformation caused by tectonic and tidal stresses. In addition, the change of the fringe width is probably related to the tilt of the tunnel wall. This method is likely to prove useful in the observation of crustal deformation.

I thank Drs Torao Tanaka, Yoshimasa Kobayashi and Kojiro Irikura for critical reading of the manuscript.

Received 31 December 1985; accepted 12 March 1986.

1. Benioff, H. *Bull. seism. Soc. Am.* **25**, 283-309 (1935).
2. Ozawa, I. *Bull. Disast. Prev. Res. Inst. Kyoto Univ.* **6**, 2-35 (1957).
3. Takada, M. *Bull. Disast. Prev. Res. Inst., Kyoto Univ.* **8**, 1-27 (1959).
4. Van Veen, H. J., Savino, J. & Alsop, L. E. *J. geophys. Res.* **71**, 5478-5479 (1966).
5. Vali, V. & Bostrom, R. C. *Rev. scient. Instrum.* **39**, 1304-1306 (1969).
6. Berger, J. & Lovberg, R. H. *Rev. scient. Instrum.* **40**, 1569-1575 (1969).
7. Levine, J. & Hall, J. L. *J. geophys. Res.* **77**, 2595-2609 (1972).
8. Gouly, N. R., King, G. C. P. & Wallard, A. J. *Geophys. J.R. astr. Soc.* **39**, 269-282 (1974).
9. Takemoto, S. *Bull. Disast. Prev. Res. Inst., Kyoto Univ.* **29**, 65-81 (1979).
10. Takemoto, S. & Kobayashi, T. *Ann. Disast. Prev. Res. Inst., Kyoto Univ.* **25**, 31-39 (1982).
11. Alyoshin, V. A., Dubrov, M. N. & Yakovlev, A. P. *Dokl. Akad. Nauk SSSR* **253**, 1343-1346 (1980).
12. Takemoto, S. *Bull. Disast. Prev. Res. Inst., Kyoto Univ.* **31**, 211-237 (1981).

The Banda-Celebes-Sulu basin: a trapped piece of Cretaceous-Eocene oceanic crust?

Chao-Shing Lee* & Robert McCabe

Geodynamics Research Program and Department of Geophysics, Texas A&M University, College Station, Texas 77843, USA

* Present address: Bureau of Mineral Resources, Geology and Geophysics, Canberra ACT 2601, Australia

The Banda, Celebes and Sulu basins are poorly understood marginal seas located at the junction of the Eurasian, Indian-Australian and Pacific-Philippine Sea plates. The incompleteness of the data sets, the complex geology of the surrounding islands and the late Cenozoic evolution involving subduction, rifting, transform faulting and island arc collision have complicated tectonic interpretation of the region. The Banda basin is underlain by oceanic crust¹⁻³. Bowin *et al.*² and Lapouille *et al.*³ have suggested that it is a trapped oceanic basin which was once continuous with the late Jurassic Argo Abyssal Plain off north-west Australia. Similarly, the Celebes and Sulu basins are also underlain by oceanic crust⁴⁻⁶. On the basis of newly identified magnetic reversal ages (Fig. 1), heat-flow data and on-land geological studies, we

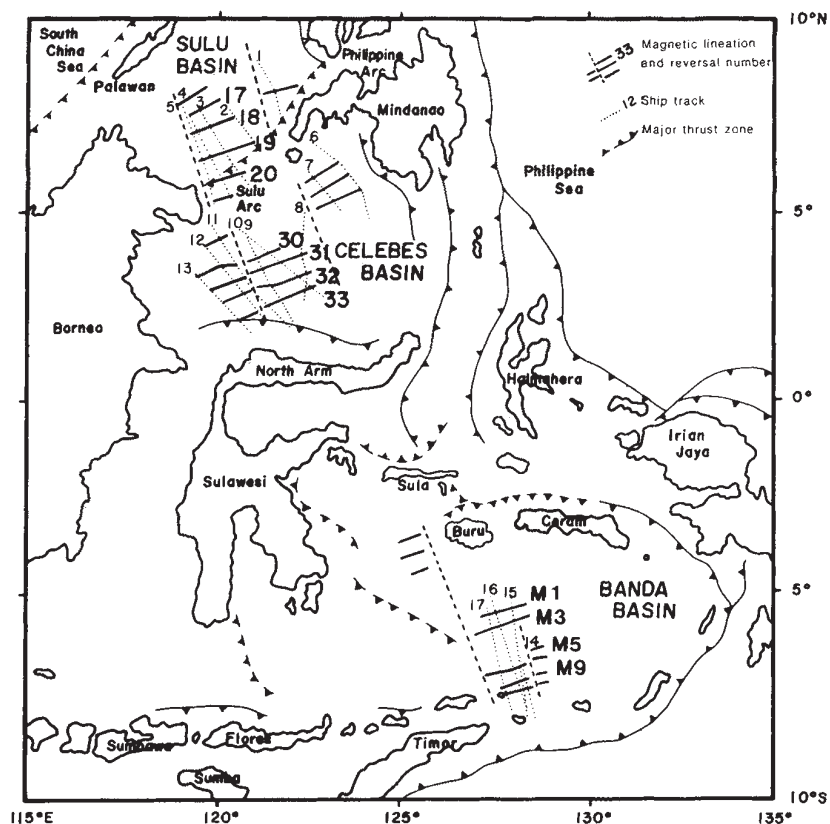


Fig. 1 Tectonic structure and magnetic anomalies in the Banda, Celebes and Sulu basins. Dotted lines are the ship tracks used for magnetic correlation in Figs 2-4. (The timescale used in this and succeeding figures is from refs 7 and 8.) These data are compiled from the global marine geophysical data set. The data sources are the Lamont-Doherty Geological Observatory (profiles 2, 4-6, 9-14 and 16), the Scripps Institute of Oceanography (1, 3, 7, 8 and 15) and the Woods Hole Oceanographic Institution (17). The tectonic structures are revised from Hamilton¹⁶.

suggest that the Celebes and Sulu seas may have been continuous with the Banda basin. This Cretaceous-Eocene basin has been dissected into its present configuration by Tertiary tectonic processes.

Re-analyses of the existing marine geophysical data suggest that the magnetic lineations in the Banda, Celebes and Sulu basins have similar trends, of about N70°E, N60°E and N55°E, respectively. Although the incomplete data sets from each individual basin do not allow magnetic ages to be determined confidently, we believe that a combined analysis of the magnetic anomalies from all three basins does permit a tentative best fit to the geomagnetic reversal model^{7,8}. Our assignments are: (1) anomalies M1-M11 (111-123 Myr BP) in the Banda basin at a half-spreading rate of $\sim 2.9 \text{ cm yr}^{-1}$; (2) anomalies 30-33 (65-72 Myr) in the Celebes basin at a half-spreading rate of $\sim 3.9 \text{ cm yr}^{-1}$; and (3) anomalies 17-20 (41-47 Myr) in the Sulu basin at 4.2 cm yr^{-1} (Figs 2-4). The magnetic anomalies in each of these three basins become younger toward the north. Similar north-east-trending late Jurassic anomalies have also been identified in the Argo Abyssal Plain⁹. These observations suggest that the Banda-Celebes-Sulu ocean basins were once part of a continuous basin, probably a northern extension of the eastern Indian Ocean¹⁰.

Recent seismic reflection and side-scan sonar studies in the Banda Sea^{11,12} have shown that the central portion of this basin is composed of prominent north-east-trending ridges and faults. Dredge data, collected in coordination with these geophysical studies, suggest that these ridges represent slivers of a continental margin. The presence of these structural features in the central portion of the Banda Sea suggests that the basin may have a complex tectonic history and is presently dissected into two or more smaller basins. This complication aside, both our model and that of Lapouille *et al.*³ for the Banda basin suggest that the entire basin is early Cretaceous in age and becomes progressively younger to the north.

Heat-flow data from the three basins are qualitatively consistent with our assigned magnetic ages: the averaged heat-flow

data fit the Parsons and Sclater¹³ heat flow versus age curve for an oceanic basin. Thus, the Banda basin² is characterized by low heat-flow values, averaging 1.09 h.f.u. ($1 \text{ h.f.u.} = 1 \mu\text{cal cm}^{-2} \text{ s}^{-1} = 41.87 \text{ mW m}^{-2}$), the Celebes basin^{6,14} by intermediate heat flow, averaging 1.38 h.f.u., and the Sulu basin¹⁴ by high heat flow; averaging 1.67 h.f.u. The fact that the heat-flow data show a good fit to the Parsons and Sclater curve also suggests that these basins are probably underlain by normal oceanic crust. (Heat flow in back-arc basins does not usually fit the global heat flow versus age curve¹⁵ this well.) The bathymetric data from the three basins also show a crude age-dependence, the Banda basin being the deepest and the basins to the north becoming progressively shallower.

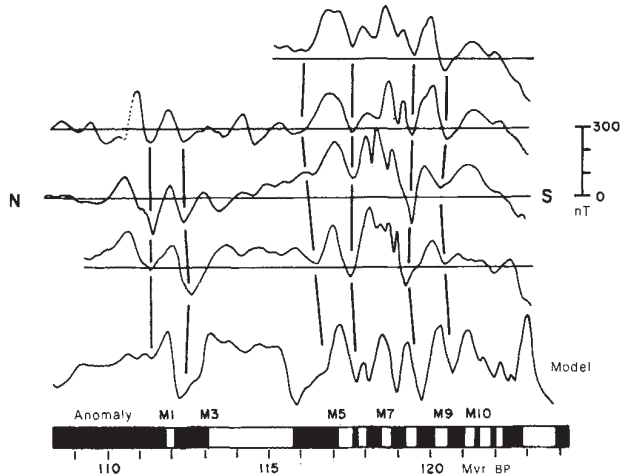


Fig. 2 Correlations of magnetic profiles with the geomagnetic timescale in the Banda basin. (The timescale used in this and succeeding figures is from refs 7 and 8.) The model is a phase-shifted synthetic magnetic profile with $\theta = -15^\circ$ where θ is the phase shift angle (see Schouten and McCamy³⁰).

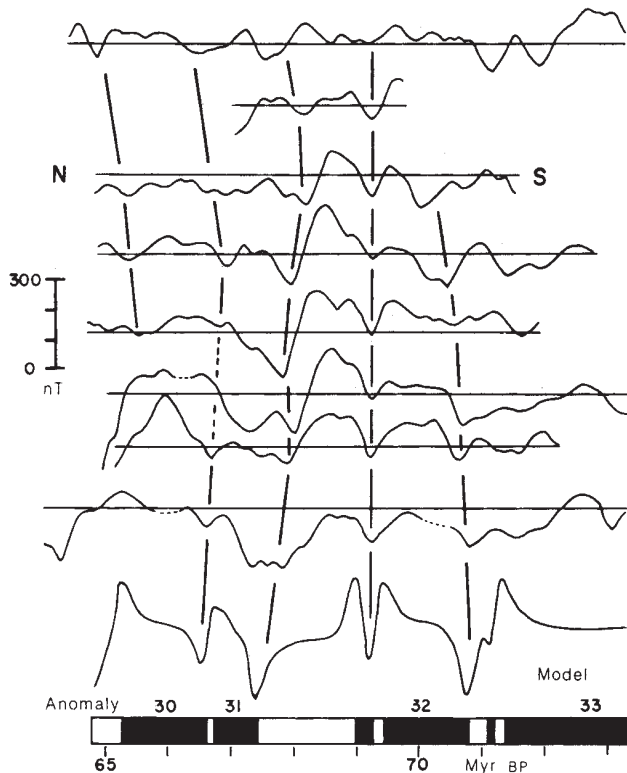


Fig. 3 Correlation of magnetic profiles with the phase-shifted synthetic model ($\theta = -330^\circ$) in the Celebes basin.

The region is complicated by numerous landmasses that divide this old oceanic basin into the present configuration of marginal seas (Fig. 1). We argue that each of these landmasses either arrived at its present location by late Tertiary tectonic movements or has been constructed in place upon this older oceanic basement. Below we briefly discuss the origin of these landmasses.

The Philippine Arc, a complex mixture of early Cretaceous to Recent island arc fragments and ophiolites¹⁶⁻¹⁹, appears to have been welded together in early to middle Tertiary time. Palaeomagnetic studies^{20,21} in the central Philippines show that this island arc has undergone motions that are consistent with the clockwise rotation of the Philippine Sea plate²². This suggests that the Philippines were located far to the south-east during the middle Tertiary. Northern Palawan is a microcontinent, composed of Palaeozoic rocks, which rifted away from southern China and translated southward during the Oligocene to Miocene opening of the South China Sea^{16,23,24}. The limited data for the Sulu Arc suggest that these islands are a late Neogene volcanic arc which developed in response to the southward subduction of the Sulu Basin¹⁶. Although the age of the island arc ophiolite is unknown, radiolarian cherts from eastern Borneo yield a late Cretaceous to Eocene age²⁵. This age is consistent with the suggestion that the basement is composed of pieces of the hypothesized Banda-Celebes-Sulu ocean basin.

Eastern Sulawesi is characterized by an ophiolite and mélange belt faulted against a metamorphosed sequence of marine strata¹⁶. Western Sulawesi, including the east-west-trending North Arm, is composed of folded sequences of Mesozoic and Tertiary metamorphic rocks, Cenozoic platform deposits and arc-derived volcanic and sedimentary rocks²⁶. Western Sulawesi probably rifted away from eastern Borneo during the middle Tertiary¹⁶. It has been suggested that the Neogene arc of the North Arm, a north-facing arc system, results from a change in subduction polarity and a clockwise rotation of the active arc^{16,27}. Although Sulawesi now divides the Celebes and Banda

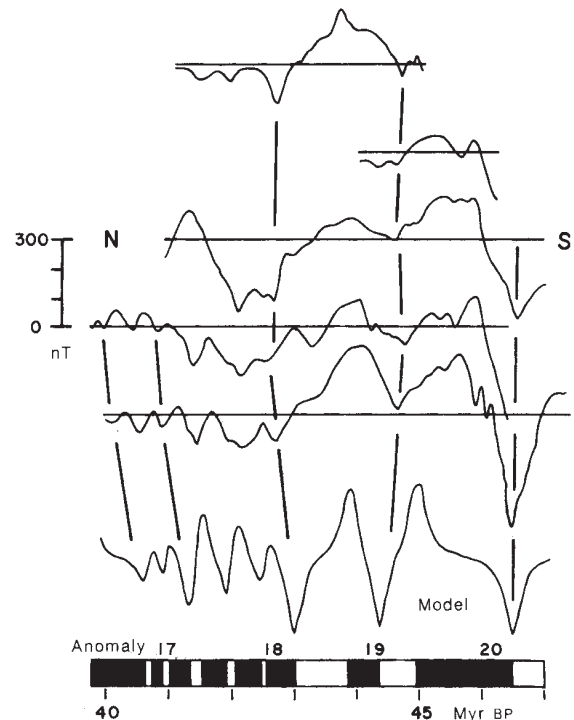


Fig. 4 Correlation of magnetic profiles with the phase-shifted synthetic model ($\theta = -270^\circ$) in the Sulu basin.

basins, we suggest that the separation of these two basins occurred during the Tertiary.

The islands of the north Banda basin, including Sula, Buru and Ceram, are composed of mélanges and continental rocks¹⁶. Stratigraphic models indicate that these islands were rifted continental fragments detached from Irian Jaya in the late Cenozoic and transported westward by strike-slip faulting²⁸. Most of the south Banda basin islands are Neogene volcanic features developed in response to late Tertiary subduction around the Banda basin¹⁶. Timor is believed to be made up of portions of the foundered Australian crust incorporated into the upper plate during the Cenozoic^{16,29}. Before this Neogene tectonics, our model suggests that this ocean and Australia were part of the same plate.

Based on marine geophysical data we propose that the Banda, Celebes and Sulu basins formed a continuous ocean basin during Cretaceous to early Tertiary times. This hypothesis is supported by the fact that the various islands which currently fragment this older basin either arrived at their present location by middle to late Tertiary tectonic movements or were built in place by Neogene subduction. The trapped portions of old oceanic crust were further reduced in size by Neogene subduction beneath the Banda Arc, the North Arm of Sulawesi and the Sulu archipelago.

We thank Tom Hilde, Robin Lighty, Jay Cole and Chris Pigram for their comments and reviews. The research was supported by the United States Department of Energy, the Center of Energy and Mineral Resources at Texas A&M University and the Amoco Production Company. This is Texas A&M Geodynamics Program Contribution No. 49.

Received 21 March; accepted 8 April 1986.

1. Purdy, G. M. & Detrick, R. S. *J. geophys. Res.* **83**, 2247-2257 (1978).
2. Bowin, C. O. *et al. Bull. Am. Ass. Petrol. Geol.* **64**, 868-915 (1980).
3. Lapouille, A. *et al. Tectonophysics* (in the press).
4. Murauchi, S. *et al. J. geophys. Res.* **78**, 3437-3447 (1973).
5. Mascle, A. & Biscarrat, P. A. *Mem. Am. Ass. Petrol. Geol.* **29**, 373-381 (1979).
6. Weissel, J. K. *Monogr. Am. geophys. Un.* **23**, 37-47 (1980).
7. Larson, R. L. & Hilde, T. W. C. *J. geophys. Res.* **80**, 2586-2594 (1975).
8. LaBrecque, J. L. *et al. Geology* **5**, 330-335 (1977).
9. Veevers, J. J. *et al. Earth planet. Sci. Lett.* **72**, 415-426 (1985).

10. Hilde, T. W. C. *et al.* *Tectonophysics* **38**, 145-165 (1977).
11. Schwartz, D., Gill, J. B. & Silver, E. A., *EOS* **64**, 872 (1983).
12. Silver, E. A., Gill, J. B., Schwartz, D., Prasetyo, H. & Duncan, R. A. *Geology* **13**, 687-691 (1985).
13. Parsons, B. & Sclater, J. G. *J. geophys. Res.* **82**, 803-827 (1977).
14. Sclater, J. G. *et al.* *J. geophys. Res.* **81**, 309-318 (1976).
15. Watanabe, T. *et al.* *Maurice Ewing Ser. Am. geophys. Un.* **1**, 137-161 (1977).
16. Hamilton, W. *Prof. Pap. U.S. geol. Surv.* No. 1078, 1-345 (1979).
17. Karig, D. E. *Tectonics* **2**, 211-236 (1983).
18. McCabe, R. *et al.* *Circum-Pacific Coun. Energy Miner. Res.* **1**, 421-436 (1985).
19. Hawkins, J. W. *et al.* *Circum-Pacific Coun. Energy Miner. Res.* **1**, 437-463 (1985).
20. Hsu, I. thesis, Washington Univ. (1972).
21. McCabe, R. *Tectonics* (in the press).
22. Loudon, K. *J. geophys. Res.* **82**, 2989-3002 (1977).
23. Taylor, B. & Hayes, D. E. *Monogr. Am. geophys. Un.* **27**, 23-56 (1983).
24. Holloway, N. H. *Bull. Am. Ass. petrol. Geol.* **66**, 1355-1383 (1981).
25. Haile, N. S. & Wang, N. P. *Mem. Malay. geol. Surv.* **16**, 1-199 (1965).
26. Sukamoto, R. & Simandjuntak, T. O. *Bull. geol. Res. Dev. Cent.* **7**, 1-12 (1983).
27. Silver, E. A. *et al.* *J. geophys. Res.* **83**, 1681-1691 (1983).
28. Visser, W. A. & Hermes, J. J. *Verh. K. Ned. Geol.-mijnb. Geroort* **20**, 1-265 (1962).
29. Audley-Charles, M. G. *Geol. Mag.* **102**, 267-272 (1968).
30. Schouten, H. & McCamy, K. *J. geophys. Res.* **77**, 7089-7099 (1972).

Seismic measurements reveal a saturated porous layer beneath an active Antarctic ice stream

D. D. Blankenship, C. R. Bentley,
S. T. Rooney & R. B. Alley

Geophysical and Polar Research Center,
University of Wisconsin-Madison,
1215 West Dayton Street, Madison, Wisconsin 53706, USA

Seismic reflection studies recently conducted on ice stream B, part of the marine ice sheet of West Antarctica, show a metres-thick layer immediately beneath the ice in which both compressional (P) and shear (S) wave speeds are very low. These low wave speeds imply that the material in the layer is highly porous and is saturated with water at a high pore pressure. From this, and from arguments presented in an accompanying paper¹ to the effect that the layer would be too weak to support the shear stress exerted by the overlying ice, we conclude that the layer is deforming and that the ice stream probably moves principally by such deformation.

For the past decade, the stability of marine ice sheets (ice sheets grounded below sea level) has stood as a fundamental unsolved problem in glaciology. In particular, the West Antarctic ice sheet has been of concern because of the possibility that it might shrink rapidly in response to climatic warming, thus raising sea level by ~5 m. But no understanding of the dynamics of the West Antarctic ice sheet is possible until the processes and interactions that govern the development and movement of ice streams, by far the most active parts of the ice sheet, are understood.

Consequently, a major program was initiated in 1983 to study the West Antarctic ice streams that feed into the Ross Ice Shelf and their interaction with that ice shelf. One aspect of this program is a concentrated geophysical study at a site on ice stream B (Fig. 1), aimed at learning more about the nature of the contact zone between the ice stream and its bed, because in this zone lies the crucial determinant of ice movement. Glaciologists agree that water lubrication must be an important factor in allowing ice streams to move so rapidly despite small driving stresses (small surface slopes), but just how the water effects the rapid sliding—whether it acts through the mechanism of a thin, approximately uniform sheet, through channels incised in the underside of the ice, or through some other mechanism—is not known²⁻⁶.

The field survey during the first season of the program (1983-84) comprised seismic recordings at ~300 sites along three lines at angles of 60° to one another, over an area of ~10 km². Geophone spreads were generally ~720 m long, with 30-m geophone spacings. All three components of motion were recorded so as to yield both P-wave and S-wave arrivals. Shots ranging

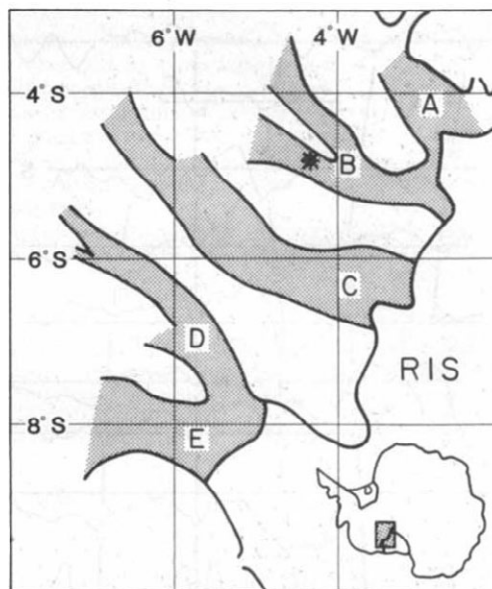


Fig. 1 Location map of ice streams along the Siple Coast of West Antarctica. An index map of Antarctica is in the lower right-hand corner. RIS, Ross Ice Shelf. Stippled areas are ice streams; the measurement site on ice stream B is marked by an asterisk. Standard grid coordinates are shown: in this rectangular system, 0° of longitude lies along the Greenwich meridian; and 0° of latitude passes through the South Pole.

in size from 0.15 to 0.45 kg were fired in shot holes ~20 m deep at spread centres and at various distances up to 4 km from the ends of the spreads. All the data were recorded on a digital recording system which was designed and built at the Geophysical and Polar Research Center. The system has a dynamic range of 84 dB and, for these experiments, sampled each of 24 channels at 0.4-ms intervals.

The existence of two reflectors near the base of the ice is seen very clearly in the seismogram of Fig. 2. In this example, the second echo is actually stronger than the first. The entire double-echo pattern is repeated 21 ms later in secondary ('ghost') seismic waves that have travelled upward from the shot to be reflected first from the surface of the glacier and then from the bed. All of the travel times in Fig. 2 have been reduced for 'normal move-out', so that reflections from a flat, level surface

Table 1 Porosities calculated by comparing v_p as measured on ice stream B with laboratory measurements

Sample	f (kHz)	Δv_p (m s ⁻¹)	Porosity	Ref.
Sand	1,000	100	0.38 ± 0.10	11
Variety of marine sediments	200	35	0.45 ± 0.18	12
Clayey silt	400	10	0.38 ± 0.10	13
Shallow water clayey silt	400	10	0.39 ± 0.09	13
Lake Erie sediments	330	0	0.41 ± 0.13	14
Artificial sands	250	100	0.33 ± 0.07	15
Unweighted mean			0.39 ± 0.06	

f is the frequency in the laboratory measurements, and Δv_p is the velocity correction relative to low frequency. Error figures on porosities correspond to a velocity uncertainty of ± 120 m s⁻¹, compounded, in an r.m.s. sense, of the uncertainty in measured velocity of ± 100 m s⁻¹, and assumed uncertainty in Δv_p of ± 50 m s⁻¹, and a characteristic r.m.s. scatter in the laboratory data of ± 40 m s⁻¹. We use an unweighted mean because we view all of the standard deviations as equally valid estimates of a single value applicable to all the measurements.

Fig. 2 Seismogram showing ice-bottom, layer-bottom, and ghost echoes, as marked. All traces have been adjusted for 'normal move-out'.

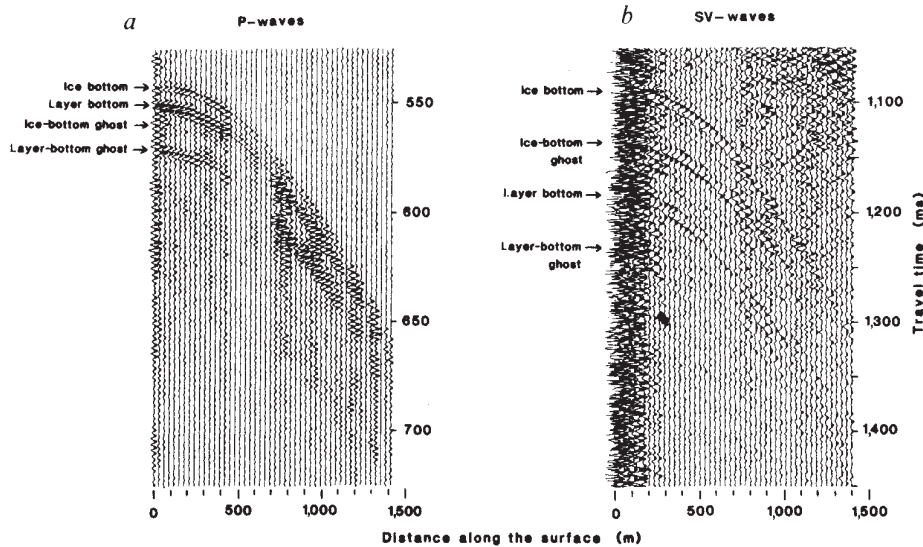
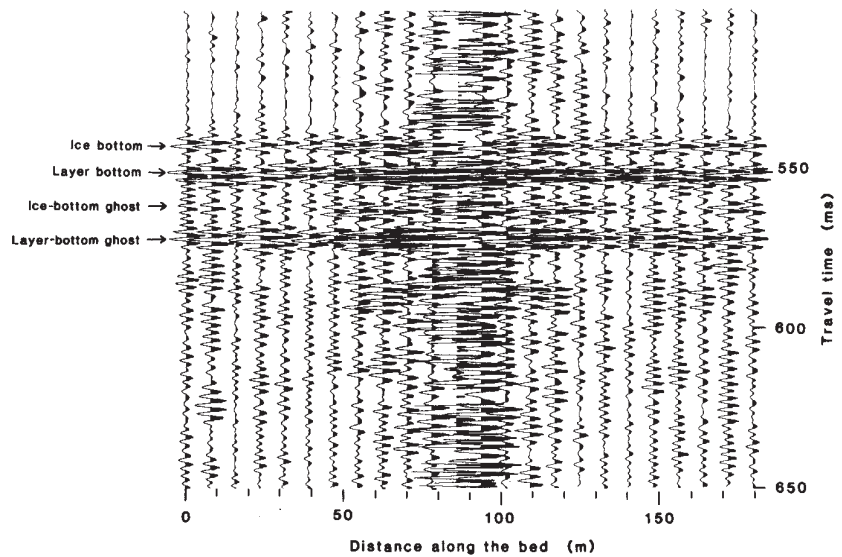


Fig. 3 Sequences of seismograms depicting 'wide-angle' (oblique) reflections. Not corrected for normal move-out. *a*, P-wave reflections; *b*, SV-wave reflections.

will appear at the same time regardless of the different horizontal distances between the shot and the various geophones.

The seismogram of Fig. 2, which was recorded along a direction parallel to the ice movement, depicts two level, coherent, reflecting surfaces separated by a travel-time difference (δt) of 10 ms. At a P-wave speed (v_p) in the layer of $1,600 \text{ m s}^{-1}$, that value of δt corresponds to a layer 8 m thick.

The wave speed in the layer was determined by means of 'wide-angle' (oblique) reflections over the section shown in Fig. 2. (The change in the travel-time difference between the two reflections as the shot and detectors are moved symmetrically outwards from the centre in opposite directions is a measure of the wave speed between the two reflectors.) The wide-angle profile (Fig. 3a) reveals that there is little change in the reflection-time difference between the two reflectors, which means that v_p in the layer is low compared with v_p in the ice ($3,830 \text{ m s}^{-1}$). Thus, we are not seeing a basal layer within the ice. Numerical inversion of the wide-angle travel times yielded $v_p = 1,600 \pm 100 \text{ m s}^{-1}$ in the layer. This precision was made possible by the large signal-to-noise ratio and the high frequency of the echoes, which permitted a timing precision of 0.2 ms. The inversion procedure fully considers the dip of both the top and bottom of the layer as well as the ray bending caused by the vertically inhomogeneous velocity structure near the surface of

the ice stream.

S-wave reflections from the upper and lower surfaces of the layer were recorded also (Fig. 3b), but with a δt that is 11 times as great as that for P-waves. This indicates an S-wave speed (v_s) of only $150 \pm 10 \text{ m s}^{-1}$ in the layer. In the Earth, v_s values this low are found only in very porous material under low effective (differential) pressure (the difference between the overburden pressure, which tends to force the grains more tightly together and thus increase the rigidity of the solid framework of the sediment, and the pore-water pressure, which tends to weaken the contact between the grains). Beneath ice stream B this must mean a saturated sediment, in which the weight of the thousand metres of overlying ice is supported principally by the pore fluid.

Along the direction of ice movement the subglacial reflecting horizons are clearly parallel to the ice bottom (Fig. 4a). The occurrence of more than one reflector suggests some stratification within the layer; the indicated strata are all nearly flat in the upstream-downstream direction. Transverse to the ice movement, however, there is much less continuity in the subglacial reflectors, and several sloping features appear (Fig. 4b). Most striking is the hump, occurring at $\sim 0.8 \text{ km}$ along the profile, that seems to rise near, or even to, the base of the ice. Since no deeper reflectors can be seen, it appears that

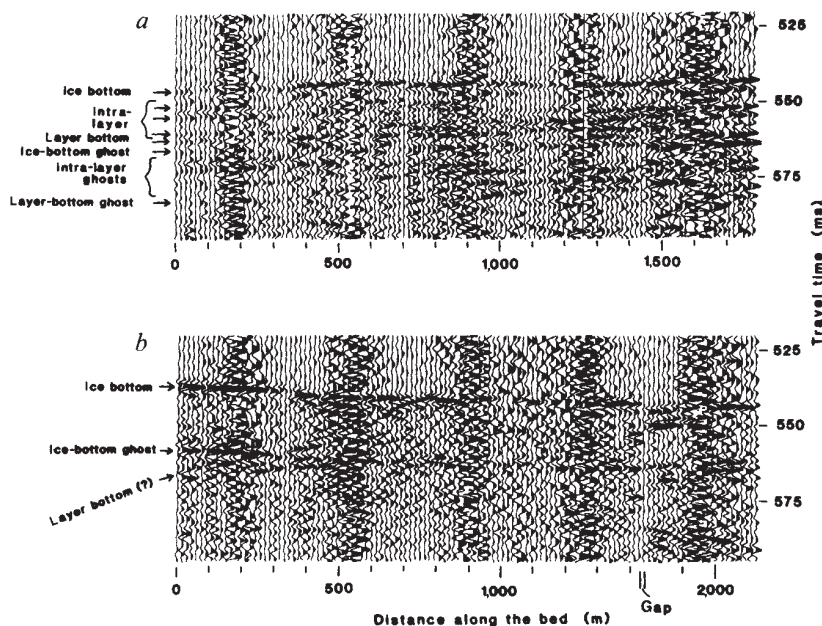
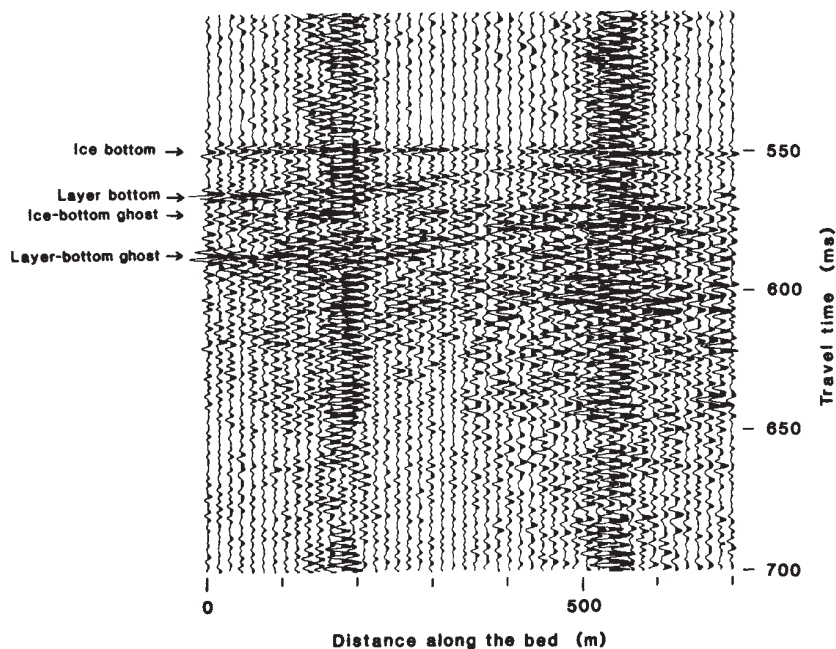


Fig. 4 Seismic reflection sections, each comprising five individual seismograms. Corrected for normal move-out. *a*, Section parallel to direction of ice movement; *b*, section transverse to ice movement.

Fig. 5 Seismic reflection section transverse to ice movement, comprising two individual seismograms. Corrected for normal move-out.



locally, the layer nearly or entirely pinches out. Another section transverse to the ice stream (Fig. 5) shows a coherent layer-bottom reflection across the entire 0.8 km with a clear topographic relief of 8 m (10 ms). The maximum layer thickness on this section is 12 m (15 ms). From the available data we estimate an average thickness of 5 or 6 m.

The picture that emerges from the longitudinal and transverse sections is one of ridges and valleys in the basal surface of the sub-glacial layer, trending parallel to the axis of the ice stream. The alignment of these features along the direction of movement suggests that they are formed by erosion. It is significant to our model of a deforming bed¹ that there is no reflection of the ridges and troughs in the glacial/sub-glacial interface.

In a porous, saturated sediment, v_p depends primarily on the porosity n (and therefore the density), whereas v_s is most sensitive to the effective pressure ΔP . We can therefore estimate both n and ΔP using the v_p and v_s we have observed for the layer. Because $v_s \ll v_p$, both the rigidity and the incompressibility of the sedimentary frame are small compared with the bulk incom-

pressibility, κ , which then can be expressed simply as $\kappa = [(1-n)\beta_s + n\beta_w]^{-1}$, wherein β_s and β_w are the compressibilities of the sedimentary particles and the water, respectively. For a typical rock particle, we take $\beta_s = 2.6 \times 10^{-11} \text{ m}^2 \text{ N}^{-1}$, and for fresh water at 0 °C, $\beta_w = 5.0 \times 10^{-10} \text{ m}^2 \text{ N}^{-1}$. From the seismic measurements, on the other hand, $\kappa = \rho v_p^2$, where the bulk sediment density $\rho = (1-n)\rho_s + n\rho_w$; ρ_s is the density of the sedimentary particles and ρ_w is the density of fresh water. We take $\rho_s = 2.6 \times 10^3 \text{ kg m}^{-3}$ and $\rho_w = 10^3 \text{ kg m}^{-3}$. Equating the two expressions for κ , we obtain $n = 0.365 \pm 0.075$ for $v_p = 1,600 \pm 100 \text{ m s}^{-1}$.

Comparisons can also be made with laboratory and *in situ* measurements of v_p versus n in saturated sediments, but two corrections are needed. First, because most of the published values of v_p refer to sediments saturated with sea water and are corrected approximately to standard laboratory temperature and pressure, we must add a correction of 100 m s^{-1} (ref. 7) to our measured velocity, which refers to fresh water at 0 °C and a glaciostatic pressure of 9 MPa.

Table 2 Effective pressures (ΔP) calculated by comparing v_s as measured on ice stream B with laboratory measurements

Sample	ΔP (kPa)	Ref.
Artificial sand	40 $\pm$ 5	15
Silt	60 $\pm$ 5	16
Silty clay	30 $\pm$ 5	16
Potter's clay	120 $\pm$ 5	16
Marine silt-clays	70 $\pm$ 50	17*
Angular-grained material	15 $\pm$ 5	17†
Sands	25	18
Clays and silts	60	18
Unweighted mean	50 $\pm$ 40	

Error figures for ΔP correspond directly to a velocity uncertainty of $\pm 10 \text{ m s}^{-1}$ except in the case of the marine silt-clays¹⁷, for which a range of observed curves was taken into account. In calculating the standard error estimate for the mean, we have assumed a standard deviation of $\pm 50 \text{ kPa}$ in ΔP for each sample.

* From Fig. 2 in ref. 17, using the supplementary relation: $\text{depth} = \Delta P / (\rho - \rho_w)g$, where g is the acceleration of gravity.

† From equation (6) in ref. 17.

Second, most laboratory measurements are made at frequencies between 100 kHz and 1 MHz, whereas our seismic frequencies are only $\sim 400 \text{ Hz}$. Standard Biot theory^{8,9} with $n = 0.4$ (to be justified below) leads to a limiting high-frequency wave speed that is $\sim 100 \text{ m s}^{-1}$ higher than the low-frequency speed. We have, therefore, added to our field measurements approximate frequency-effect corrections (Δv_p in Table 1) calculated according to standard theory¹⁰. These corrections are in approximate agreement with those measured for 1 MHz for various sediments⁹, but because Δv_p depends on permeabilities in the sediment samples that had to be estimated, we assume an uncertainty in Δv_p of $\pm 50 \text{ m s}^{-1}$. Fuller details will be given elsewhere.

The porosities obtained from these comparisons show considerable scatter (Table 1) but are in satisfactory agreement with the value calculated above by linearly combining compressibilities and densities. It therefore seems safe to conclude that n in the sub-glacial layer is close to 0.4, although probably a bit less. That is is close to 0.4 and not 0.3 or less is significant, as it strongly suggests that the sub-glacial material is dilated and deforming¹.

From v_p and ρ in the layer we can calculate that the acoustic impedance of the sediment is close to that of the ice. This explains how the echo from the bottom of the sediment could be stronger than that from the base of the ice. Echo amplitudes vary widely over the survey area, however; they will be considered in detail elsewhere.

The other quantity we wish to estimate is ΔP , since it is critical to the shear strength of the medium. The best measure of ΔP is v_s , because in an unconsolidated sediment v_s depends principally on the intergranular friction, and therefore on the intergranular pressure. Again, comparison with several experimental values leads to a considerable uncertainty (Table 2); nevertheless, all estimates of ΔP lie in the range 15–120 kPa. We adopt the mean value of $50 \pm 40 \text{ kPa}$ for ΔP in the sub-glacial layer.

We conclude that at least at one location, an active Antarctic ice stream is underlain by a layer of saturated sediment ranging in thickness from zero or nearly zero to at least 12 m, with an average of 5 or 6 m. The layer varies much less in thickness parallel to the direction of ice movement than normal to it; because its upper surface is planar, this gives the impression of a series of longitudinal grooves in the substrate beneath the layer. The low seismic wave speeds in the layer indicate that the material has a porosity of ~ 0.4 and is saturated with water at a pore pressure only $\sim 50 \text{ kPa}$ less than the glaciostatic pressure (9 MPa). Because these characteristics imply that the sub-glacial material is very weak¹, we believe that the sub-glacial

layer is deforming and eroding the stationary surface below, and that it is deformation within the layer rather than deformation in the ice or basal sliding that is the principal component of ice-stream movement.

The field program is a cooperative effort among investigators from the University of Wisconsin, the Ohio State University, NASA and the University of Chicago. We thank B. R. Weertman, J. E. Nyquist, R. Flanders, S. Shabtaie, D. G. Schultz, K. C. Taylor, Jr and J. Dallman for help with field work, L. A. Powell, B. D. Karsh, R. B. Abernathy, W. L. Unger, and Weertman for engineering assistance, and A. N. Mares and S. H. Smith for manuscript and figure preparation. Financial support was provided by NSF under grants DPP81-20322 and DPP84-12404. This is contribution number 447 of the University of Wisconsin-Madison, Geophysical and Polar Research Center.

Received 23 December 1985; accepted 15 April 1986.

- Alley, R. B., Blankenship, D. D., Bentley, C. R. & Rooney, S. T. *Nature* **322**, 57–59 (1986).
- Iken, A. *J. Glaciol.* **27**, 407–421 (1981).
- Weertman, J. & Birchfield, G. E. *Ann. Glaciol.* **3**, 316–320 (1982).
- Bindschadler, R. J. *J. Glaciol.* **29**, 3–19 (1983).
- Iken, A., Rothlisberger, H., Flotron, A. & Haeberli, W. *J. Glaciol.* **29**, 28–47 (1983).
- Kamb, B. *et al. Science* **227**, 469–479 (1985).
- Clay, C. S. & Medwin, H. *Acoustical Oceanography: Principles and Applications* (Wiley, New York, 1977).
- Biot, M. A. *J. acoust. Soc. Am.* **28**, 168–191 (1956).
- Hamdi, F. A. I. & Taylor Smith, D. *Geophys. Prospect.* **30**, 622–640 (1982).
- Geertsma, J. & Smith, D. C. *Geophysics* **26**, 169–181 (1961).
- Hamdi, F. A. I. & Taylor Smith, D. *Geophys. Prospect.* **29**, 715–729 (1981).
- Hamilton, E. L. *J. geophys. Res.* **75**, 4423–4445 (1970).
- Anderson, R. S. in *Physics of Sound in Marine Sediments* (ed. Hampton, L.), 481–518 (Plenum, New York, 1974).
- Morgan, N. A. *Geophysics* **34**, 529–545 (1969).
- Schultheiss, P. J. in *Acoustics and the Sea Bed* (ed. Pace, N. G.) 19–27 (Bath University Press, 1983).
- Schultheiss, P. J. *Marine Geotechnol.* **4**, 343–367 (1981).
- Hamilton, E. L. *Geophysics* **41**, 985–996 (1976).
- Ohta, V. & Goto, N. *Earthq. Engng struct. Dyn.* **6**, 167–187 (1978).

Deformation of till beneath ice stream B, West Antarctica

R. B. Alley, D. D. Blankenship,
C. R. Bentley & S. T. Rooney

Geophysical and Polar Research Center, University of Wisconsin-Madison, 1215 West Dayton Street, Madison, Wisconsin 53706, USA

The behaviour and possible instability of the West Antarctic ice sheet depend fundamentally on the dynamics of the large ice streams which drain it. Model calculations show that most ice-stream velocity arises at the bed^{1,2}, and radar sounding has shown the bed to be wet³, but the basal boundary condition is not well understood. Seismic evidence from the Upstream B camp (UpB) on the Siple Coast of West Antarctica⁴ shows that the ice stream there rests on a layer of unconsolidated sediment averaging 5 or 6 m thick, in which the water pressure is only $\sim 50 \text{ kPa}$ less than the overburden pressure. Because this thin layer occurs well inland beneath an active ice sheet and rests on a surface showing flutes⁴ characteristic of glacial erosion⁵, we presume that it is glacial till. We propose here that deformation within the till is the primary mechanism by which the ice stream moves, and we discuss implications of this hypothesis.

Three distinct lines of evidence relating to till porosity, force balance, and water balance indicate that the till at UpB is deforming; we present each of these here. First, lodged basal till has a porosity of $\approx 30\%$, whereas deformation of till causes dilation and porosity of $\sim 40\%$ (refs 6–8). The seismically determined porosity of $\sim 40\%$ for till at UpB⁴ is consistent with active deformation but too high for lodged till.

Second, we estimate that the basal shear stress is about twice the strength of till at UpB, so that the till should be deforming.

The strength of till can be estimated from measured properties of deforming till, collected by Boulton and co-workers⁶⁻⁹ beneath Breidamerkurjökull in Iceland. Briefly, Boulton found that the basal-ice velocity of Breidamerkurjökull arises largely from deformation of a dilated till layer of 40% porosity, that large stress concentrations occur locally during deformation, and that the dilated structure collapses and till strength increases by a factor of 1.6–4 if the strain rate is too low.

Because of the large stress concentrations in deforming till, ice-stream flow would dilate and deform lodged, low-porosity till if the average basal shear stress, τ , exceeded the strength of dilated till, τ_d . This is because, if the till at UpB were lodged, then the basal-ice velocity would arise from sliding across a water layer that reduced ice-bed contact to local patches. Weertman¹⁰ has shown that total ice-bed contact area would have to be reduced greatly to allow sliding at ice-stream velocities, so that the local shear stress in patches of contact would significantly exceed 4τ . However, the strength of lodged till is $<4\tau_d$ (refs 6–8), so that local mobilization of till would occur for $\tau \geq \tau_d$. The stress concentrations caused by deformation would then mobilize adjacent, lodged till to form a continuous deforming layer that would thicken to intersect bedrock or until the strain rate became too small to generate stress concentrations large enough to mobilize more till. (Till would also be mobilized locally for some range of $\tau < \tau_d$, but the average basal shear stress would not be large enough in such conditions to deform a continuous layer of till.)

Boulton⁹ argues that till obeys a Coulomb-type failure criterion, so we let $\tau_d = \Delta P \tan \phi_d$, where ΔP is the excess of overburden pressure over water pressure and $\tan \phi_d$ is the internal friction. (We have set cohesion to zero in the failure criterion, in accordance with the observations of Engelhardt *et al.*¹¹ and because we expect dilation of till to disrupt the short-range electrostatic forces between clay particles that cause the cohesion in saturated, lodged till.) Till is deforming at Breidamerkurjökull, so τ must be $\geq \tau_d$ there. Using measured ΔP (ref. 7) and basal shear stress calculated from the geometry of Breidamerkurjökull⁹, we estimate from the Coulomb failure criterion that $\tan \phi_d \leq 0.15$. (Direct measurements of recently deformed till at Breidamerkurjökull⁶⁻⁸ using a shear box give values of $\tan \phi_d$ equal to or slightly larger than that reported here, but laboratory shear tests consistently overestimate the strength of materials undergoing slow, large-scale deformation¹².) At UpB, $\Delta P = 50 \pm 40$ kPa (ref. 4); if internal friction of till is similar at UpB and Breidamerkurjökull, then $\tau_d \leq 8 \pm 6$ kPa at UpB.

The driving stress for ice flow at UpB is ~ 19 kPa, calculated from the familiar formula: driving stress $= \rho gh \tan \alpha$, where the ice density $\rho = 920$ kg m⁻³; the gravitational acceleration $g = 9.8$ m s⁻²; the ice thickness $h = 1,050$ m, and the surface slope $\tan \alpha \approx 0.002$ (ref. 13). The driving stress is resisted by gradients in stretching stress and by side drag as well as by the basal drag, τ , but a force-balance calculation for the ice stream by Whillans¹⁴ indicates that basal drag is the major resistive force. We have reached the same conclusion independently, and estimate that $\sim 95\%$ of the driving stress is balanced by basal drag at UpB, so that $\tau = 18$ kPa. This is more than twice the best estimated value of τ_d at UpB, so we expect till deformation to be occurring there.

The third argument for till deformation comes from a water-balance calculation. As noted above, if till is not deforming then ice-stream velocity must arise from sliding of ice over a rigid substrate lubricated by water. Weertman¹⁰ estimates that fast sliding requires a water film of thickness $d \approx 5$ mm or greater. At the head of an ice stream fed by a catchment area of length $L \approx 400$ km and width $F \approx 4$ times as wide as the ice stream¹³, the existence of fast sliding requires production averaged over the catchment area of a thickness of water per unit time, λ , given by¹⁰: $\lambda = d^3 \rho g \tan \alpha (12 \eta L F)^{-1}$, where $\rho g \tan \alpha$ is the pressure gradient driving water flow in the ice stream (assuming bed slope is of the same magnitude as surface slope or smaller¹³)

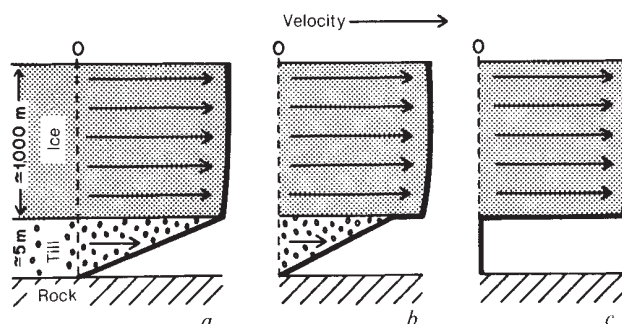


Fig. 1 Possible models for an ice-stream bed. *a*, Till deformation only; our model for UpB. *b*, Till deformation plus basal sliding; our model for near the grounding line of line stream B. *c*, Basal sliding only; not realized on ice stream B.

and $\eta = 1.8 \times 10^{-3}$ Pa s is the viscosity of water¹⁰. The minimum water production for fast sliding is then $\lambda \approx 2$ mm yr⁻¹. We consider this much water production to be unlikely because: (1) the Byrd borehole, in the catchment area of ice stream D, showed that net basal freezing rather than melting occurs over some distance up-glacier of the borehole¹⁵, and the catchment of ice stream B is sufficiently similar to that of ice stream D in surface mass balance, surface topography, bed topography and ice thickness¹³ that we expect net basal freezing in parts of the ice-stream-B catchment; (2) in numerical experiments, Budd *et al.*¹⁶ were able to generate a melt rate of 2 mm yr⁻¹ for most of the catchment only by increasing the geothermal heat flow by 50% above the value they expected based on the geology of West Antarctica; their expected value has been confirmed by measurements in the Byrd Station borehole^{17,18}; and (3) ice stream B and its catchment are underlain by ≥ 1 km of sedimentary rock^{19,20} that may have sufficient permeability²¹ to drain a significant amount of water. The amount of water available to lubricate sliding would be reduced further if some water flowed in channels rather than in a Weertman film²². It is thus unlikely that enough water is supplied to the ice stream to allow high velocities near its head by sliding. (Note, however, that water generated beneath the ice stream may form a film at the ice-till interface²³, so that both till deformation and water-lubricated sliding may be important near the grounding line (Fig. 1).)

The above arguments show that most ice-stream velocity arises from till deformation, and we now use this result to show that the entire thickness of till at UpB is deforming. As noted above, stress concentrations in active till increase with increasing strain rate, $\dot{\epsilon}$. If $\dot{\epsilon}$ were to exceed $\dot{\epsilon}_b$, the minimum strain rate required to cause stress concentrations large enough to mobilize lodged till, then active till would mobilize subjacent, lodged till until the strain rate was reduced to $\dot{\epsilon}_l$ or until till activation reached bedrock. We thus expect $\dot{\epsilon} = \dot{\epsilon}_l$ if active till overlies lodged till and $\dot{\epsilon} \geq \dot{\epsilon}_l$ if active till reaches bedrock. The average simple strain rate in till is u/h_b , where u is the velocity of the till at its upper surface and h_b is the thickness of the active layer. At Breidamerkurjökull⁹, where active till overlies lodged till, $u \approx 16$ m yr⁻¹, $h_b \approx 0.5$ m and $\dot{\epsilon} \approx 32$ yr⁻¹ $= \dot{\epsilon}_b$. At Blue Glacier¹¹, where active till reaches bedrock, $u \approx 4$ m yr⁻¹, $h_b \approx 0.1$ m and $\dot{\epsilon} \approx 40$ yr⁻¹ $\geq \dot{\epsilon}_b$. At UpB, for little slip between ice and till, $u \approx 450$ m yr⁻¹ (ref. 24) and $h_b \approx 6$ m, so $\dot{\epsilon} \approx 75$ yr⁻¹. As this is much greater than $\dot{\epsilon}_l$ at Blue Glacier and Breidamerkurjökull, and there is no reason for $\dot{\epsilon}_l$ to be grossly different at UpB, we conclude that the entire thickness of till at UpB is deforming.

Given that the entire thickness of till at UpB is deforming, we can now calculate the fluxes of till and water beneath UpB and the generation rates in the catchment area required to sustain these fluxes if they are steady-state values. Water moves by advection with till and by conduction through till. From Darcy's law, the conductive flow velocity is $u_c = k/(\eta(dP/dx))$ where

k = till permeability $\approx 1.6 \times 10^{-13} \text{ m}^2$ (ref. 7), η -dynamic viscosity of water $= 1.8 \times 10^{-3} \text{ Pa s}$, and (dP/dx) = pressure gradient from weight of ice $= \rho g \tan \alpha \approx 18 \text{ Pa m}^{-1}$ (again assuming that bed slope is of the same magnitude as surface slope or smaller¹³). The flow velocity, u_c , is then $1.6 \times 10^{-9} \text{ m s}^{-1} \approx 0.05 \text{ m yr}^{-1}$. The depth-averaged advective flow velocity, $\bar{u}_a$, is given by $\bar{u}_a = \bar{u}n$, where $\bar{u}$ is the depth-averaged velocity of bulk till and $n = 0.4$ is the till porosity⁴. For no-slip boundaries and any reasonable till rheology (such as linear-viscous, power-law creep with finite exponent, or Bingham substance) velocity will vary almost linearly from 450 m yr^{-1} at the ice-till interface²⁴ to zero at the till-bedrock interface, so $\bar{u} \approx 225 \text{ m yr}^{-1}$ and $\bar{u}_a \approx 90 \text{ m yr}^{-1}$. This is more than three orders of magnitude larger than the conductive velocity, so essentially all the water flow is advective.

The total water flux per unit width of ice stream is $h_b \bar{u}_a w$, where $h_b \approx 6 \text{ m}$ is the till thickness and w is the ice-stream width. In steady state, this flux corresponds to an average water-generation rate of thickness λ over an area $(FL + L_2)w$, where $L_2 \approx 100 \text{ km}$ is the length of the ice stream above UpB, $L \approx 400 \text{ km}$ is the length of the catchment area, and $F \approx 4$ is the width of the catchment area feeding unit width of the ice stream¹³. Then $\lambda = h_b \bar{u}_a (LF + L_2)^{-1} \approx 0.3 \text{ mm yr}^{-1}$. This is almost an order of magnitude less than the 2 mm yr^{-1} melt rate needed for Weertman sliding at the head of the ice stream, and is consistent with our knowledge of water generation in the catchment area. The average erosion rate of rock can be obtained from the same calculation by replacing porosity, n , by rock fraction, $1-n$, and is $\sim 0.5 \text{ mm yr}^{-1}$.

We thus hypothesize that the velocity of ice stream B near UpB arises largely from deformation of a sub-glacial till layer (Fig. 1). This hypothesis suggests several corollaries, which include: (1) the basal boundary condition for ice-stream flow depends on the constitutive relation for till; (2) the till flux beneath UpB is equivalent to steady-state erosion of $\sim 0.5 \text{ mm yr}^{-1}$ of rock over the catchment area and the upstream part of the ice stream; (3) this till flux requires deposition of morainal banks or 'till deltas' at the grounding line; recent mapping of the grounding line²⁵ is consistent with the existence of such deltas; (4) the fluted nature of the till-rock interface at UpB may be a characteristic erosional consequence of deforming till²⁶; and (5) variations in till thickness over flutes may cause fluctuations in basal drag that affect ice-stream flow, with higher drag over bedrock ridges and lower drag over troughs. Our work also suggests the possibility that other ice streams and other wet-based regions of ice sheets²⁷ may rest on till. We are conducting further field work and analysis to test these hypotheses.

This work was supported in part by NSF under grants DPP-8120322 and DPP-8412404. We thank I. M. Whillans for helpful comments on the manuscript, and A. N. Mares and S. H. Smith for manuscript and figure preparation. This is contribution number 448 of the Geophysical and Polar Research Center, University of Wisconsin-Madison.

Received 24 December 1985; accepted 21 April 1986.

1. Alley, R. B. *Inst. Polar Stud. Rep. No. 84* (The Ohio State University, 1984).
2. Budd, W. F., Janssen, D. & Smith, I. N. *Ann. Glaciol.* **5**, 29-36 (1984).
3. Robin, G. de Q., Swinbank, C. W. M. & Smith, B. M. E. *JASH Publ.* **86**, 97-115 (1970).
4. Blankenship, D. D., Bentley, C. R., Rooney, S. T. & Alley, R. B. *Nature* (this issue).
5. Gravenor, C. P. & Menzies, W. A. *Am. J. Sci.* **256**, 715-728 (1958).
6. Boulton, G. S. & Dent, D. L. *Geogr. Annl.* **A56**, 121-134 (1974).
7. Boulton, G. S., Dent, D. L. & Morris, E. M. *Geogr. Annl.* **A56**, 135-145 (1974).
8. Boulton, G. S. & Paul, M. A. *Q. J. Engng Geol.* **9**, 159-194 (1976).
9. Boulton, G. S. *J. Glaciol.* **23**, 15-37 (1979).
10. Weertman, J. *Can. J. Earth Sci.* **6**, 929-939 (1969).
11. Engelhardt, H. F., Harrison, W. D. & Kamb, B. J. *Glaciol.* **20**, 469-508 (1978).
12. Walker, B. F. in *In-situ Testing for Geotechnical Investigations* (ed. Ervin, M. C.) 65-72 (Balkema, Rotterdam, 1983).
13. Drewry, D. J. (ed.) *Antarctica: Glaciological and Geophysical Folio* (Scott Polar Research Institute, Cambridge, 1983).
14. Whillans, I. M. *Proc. Int. Workshop, Stability of West Antarctic Ice Sheet and Sea Level* (Reidel, Dordrecht, in the press).
15. Gow, A. J., Epstein, S. & Sheehy, W. J. *Glaciol.* **89**, 185-192 (1979).
16. Budd, W. F., Janssen, D. & Radok, U. *ANARE Interim Rep. Ser. A(IV)* (Glaciology) No. 120 (1971).

17. Gow, A. J., Ueda, H. T. & Garfield, D. E. *Science* **161**, 1011-1013 (1968).
18. Rose, K. E. *J. Glaciol.* **24**, 63-74 (1979).
19. Bentley, C. R. & Clough, J. W. in *Antarctic Geology and Geophysics* (ed. Adie, R. J.) 683-691 (Universitetsforlaget, Oslo, 1972).
20. Rooney, S. T., Blankenship, D. D. & Bentley, C. R. in *Proc. 6th Gondwana Symp.* (American Geophysical Union, in the press).
21. Freeze, R. A. & Cherry, J. A. *Groundwater* (Prentice-Hall, New Jersey, 1979).
22. Bindschadler, R. A. *J. Glaciol.* **29**, 3-19 (1983).
23. Weertman, J. & Birchfield, G. E. *Ann. Glaciol.* **3**, 316-320 (1982).
24. Vornberger, P. O. & Whillans, I. M. *Ann. Glaciol.* **7**, (in the press).
25. Shabtaie, S. & Bentley, C. R. *Ann. Glaciol.* **8**, (in the press).
26. Also see MacClintock, P. & Dreimanis, A. *Am. J. Sci.* **262**, 133-142 (1964).
27. Oswald, G. K. A. & Robin, G. de Q. *Nature* **245**, 251-254 (1973).

Corals on seamount peaks provide evidence of current acceleration over deep-sea topography

Amatzia Genin, Paul K. Dayton,
Peter F. Lonsdale & Fred N. Spiess

Scripps Institution of Oceanography, University of California,
San Diego, La Jolla, California 92093, USA

Geological and physical studies of seamounts have suggested the existence of distinct deep-sea habitats, characterized by exposed rocky bottom and a unique current regime¹⁻⁹. However, few biological data have been collected for deep seamounts¹⁰⁻¹². Here we present some of the first quantitative observations of hard-bottom (non-hydrothermal) fauna in the deep sea. These observations show that black corals (antipatharians) and horny corals (gorgonians) present on the slopes of a multi-peaked seamount are more abundant near peaks, compared with mid-slope sites at corresponding depths. On narrow peaks corals are most abundant on the crest, whereas on wide peaks, coral densities are highest at the edge of the crest. The abundance of corals also increases on knobs and pinnacles. Physical models and observations^{2,4-9,13-15}, together with our direct measurements, suggest that the seamount topography affects the local current regime. Corals appear to benefit from flow acceleration, and some of their patterns of distribution can be explained by current conditions. These results suggest that suspension feeders have some potential as indicators of prevailing currents at deep hard-bottom sites.

We made our observations on Jasper Seamount, which is an extinct volcano 3.5 km high 30 km in diameter, built on the 4,200-m-deep abyssal Pacific floor, about 550 km south-west of San Diego, California. It has a multi-peaked summit area (Fig. 1) with two of the peaks rising higher than 600 m depth. Most of the summit area consists of bare basaltic rock, with thin pockets of sediment in some depressions between the peaks. A rich and diverse fauna, dominated by sedentary suspension feeders such as sponges, black corals, horny corals, anemones and tunicates, was discovered in our photographic survey of the seamount. The most common species on the upper slopes of the seamount is the spiral black coral, *Stichopathes* sp., which occupies rocky substrata at depths from 550 m to 1,150 m, where it sometimes forms 'forests' with densities of up to 20 colonies per m². In areas partially covered with sediment, *Stichopathes* is found only on protruding rocks; it is absent in areas with 100% sediment cover. Along its upper range, the *Stichopathes* zone overlaps a rich sponge zone. Gorgonians and branched antipatharians are the dominant taxa on the rocky bottom at depths below the *Stichopathes* zone. However, coral densities at these greater depths are much lower than in the *Stichopathes* zone.

Variations of coral abundances, at any depth, are associated with the local topography, on three different scales:

(1) Over the entire seamount, *Stichopathes* densities are significantly higher near peaks than at mid-slope sites at corresponding depths (Fig. 2). For example, the average *Stichopathes*

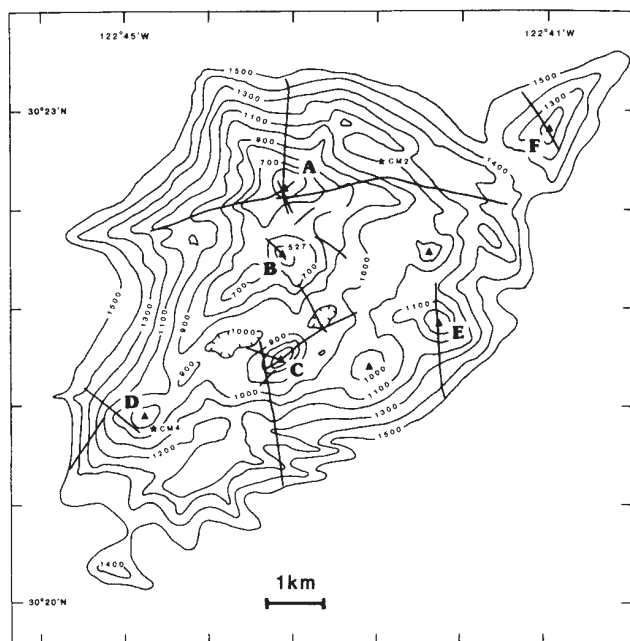


Fig. 1 The summit of Jasper Seamount, showing bathymetry, designated peak names, locations of current-meter moorings (★), and photographic transect lines (heavy black lines). Contours are in metres. Photographs were taken by the Deep-Tow instrument package²⁰ at ~10-m intervals. The camera-to-bottom distance was recorded when each photograph was taken, allowing computation of areas and animal densities.

density near peak C, between 700 and 750 m depth, is about 15 colonies per m² (Fig. 3), which is about three times the mean density of the species on slopes at the same depth interval (Fig. 2). Similarly, densities of gorgonians near peak F, at a depth of 1,170–1,250 m (Fig. 3), are more than five times higher than at the same depth interval on the slopes.

(2) On the peaks, the species exhibit two distribution patterns: on wide peaks (diameter > 1 km), with a flat or gradually sloping top, highest densities are found along the perimeter of the top, and lower densities on the top itself (Fig. 3, peaks A and F); on narrow, tapering peaks, highest densities are found on the crests (Fig. 3, peak C).

(3) On a smaller scale, *Stichopathes* exhibits increased abundances on knobs and pinnacles. For example, a sharp increase in density is found on the small knobs at 1.4 km and 3.9 km along the transect across peak A (Fig. 3).

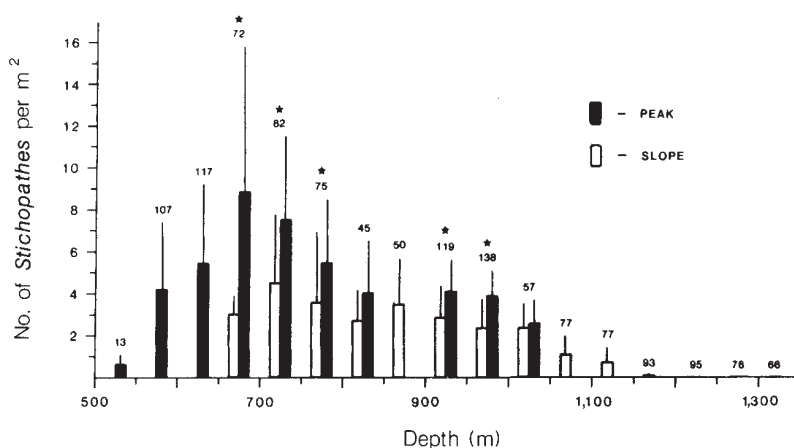


Fig. 2 Densities of *Stichopathes* (mean ± s.d.) at different depth intervals on Jasper Seamount. The photographs were divided into those near peaks (solid bars) and those >150 m below a peak (open bars). Only photographs which exhibited <10% sediment cover were included in the computation. Stars indicate the intervals in which the mean near-peak density is significantly higher than the mean mid-slope density ($P < 0.001$, Mann-Whitney U-test). The number of photographs taken at each depth interval is indicated above the corresponding bar.

Why are corals more abundant near peaks? We assume that concentrations of food and larvae in waters impinging on the seamount are independent of the local topography, that is, the average concentrations of food and larvae in waters impinging on a peak and on a slope at the same depth are similar. However, fluxes of food and larvae could vary at different sites due to topographic effects on water flow. Within a certain range of current speed, areas of flow acceleration are expected to be favourable for recruitment and growth of passive suspension feeders; this may be attributed to either the 'settlement pathway', in which a site is colonized by relatively more recruits simply because more water, and hence more larvae, are flowing by per unit of time, or the 'feeding pathway', in which an increase of water flow past suspension feeders results in increased feeding and growth rates¹⁶⁻¹⁸. This may also be critical for the survival of small recruits. Both pathways can lead to long-term integration of current conditions by the animals.

There are two lines of evidence for intensified water flow near peaks, where the abundance of corals is relatively high. First, our direct current measurements on Jasper Seamount (Fig. 4), although short-term and limited to two stations, showed that the mean total speed near a peak was about twice the mean speed at the same depth but on a mid-slope site. Second, observations on other seamounts^{9,13-15} have demonstrated hydrographic conditions near the summit which differ from those along the slopes, because the upwelling requires acceleration over the top of the seamounts.

The distribution of corals on wide peaks can be explained by predictions of physical theory. When water is upwelled above wide topography, vortex lines are compressed and anticyclonic motion is induced, because of conservation of potential vorticity^{4,5}. The resulting flow-field is a combination of the velocity induced by this anticyclonic component and the free-stream velocity. Thus, looking downstream, flow acceleration occurs on the left side of the topographic feature (in the Northern Hemisphere) and deceleration above the centre and right side of the peak^{3,4,7}. As this theory has been developed for a uniform and steady flow, the validity of the model for strong tidal currents (see, for example, Fig. 4) is less certain. A possible effect of the tide would be a switch of the anticyclonic acceleration-deceleration zones. The overall mean velocity pattern would therefore comprise alternating periods of acceleration on the peak's edges and recurrent deceleration trend over the centre. This postulated velocity pattern is consistent with the lower densities of corals observed at the centre of the wide peaks of Jasper Seamount (Fig. 3). Furthermore, the combination of tidal fluctuations and a general southward flow in the vicinity of the seamount (Fig. 4; ref. 19) could result in a stronger average acceleration on the eastern side of the seamount; this could explain why *Stichopathes* is about twice as dense on the eastern edge of the main seamount peak as on the western edge (Fig. 3, peak A).

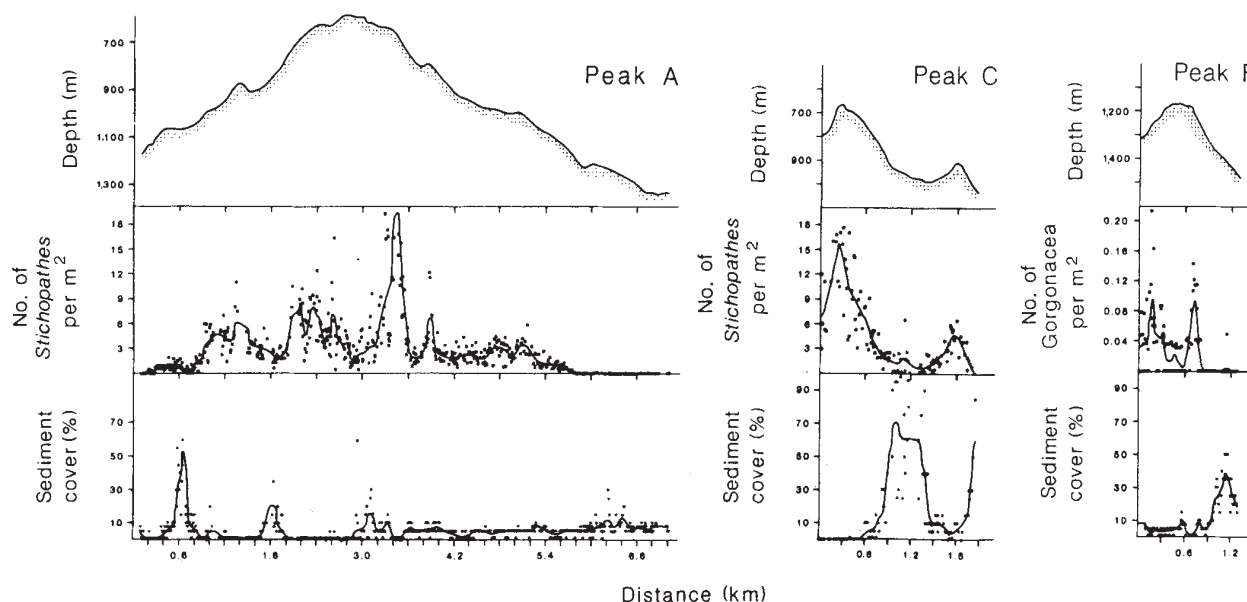
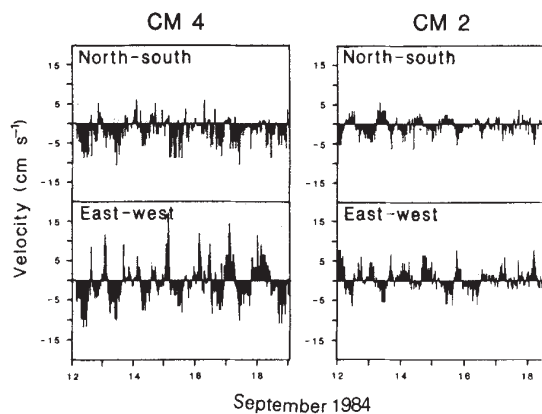


Fig. 3 Bathymetry, animal densities and sediment cover in transects across peaks A, C and F. Each dot indicates the animal density or sediment cover in a single photograph. Solid lines indicate 9-point running means. The transects are from west to east across peaks A and C, and from south-east to north-west across peak F. Sediment cover, which can have a deleterious effect on suspension feeders^{21,22}, cannot explain why corals are more abundant near peaks and on pinnacles. For example, the distribution pattern exhibited by *Stichopathes* close to peak C occurs along the initial 600 m of the transect, which is sediment-free. Similarly, distinct changes in gorgonian densities near peak F occur in an area where the sediment cover is low and nearly uniform. Overall, the correlation between sediment cover and *Stichopathes* density between 527 and 1,150 m depth is low but significantly different from zero (Kendall rank correlation = -0.2 , $P < 0.01$, $n = 1,359$).

Fig. 4 (Right) Current velocity and progressive vector diagram of records taken near peak D (CM 4) and at a mid-slope site east of peak A (CM 2), both at a depth of $\sim 1,000$ m. Dots on the progressive vector diagram of CM 4 indicate 24-h intervals. The mean current speed (15-min averages) was significantly higher ($P < 0.001$, Student's *t*-test) near the peak (4.9 cm s^{-1}) than at the mid-slope site (2.7 cm s^{-1}). Measurements were made with two SIO-Model 6 Savonius rotor current meters, deployed at 46 m above the bottom. Spectral analysis of these records indicates that the kinetic energy of internal waves near a critical frequency ($0.5 \text{ cycles h}^{-1}$), in which internal waves are reflected parallel to the topographic slope^{8,23,24}, is negligible compared with the energetically dominant diurnal frequency. Such reflection, when present, may intensify water flow over seamounts and slopes⁸.



We conclude that the current regime is a key factor in determining community structure on the deep hard-bottom areas of Jasper Seamount, where suspension feeders are more abundant at sites of flow acceleration.

We thank D. Boegman, R. Elder, D. Genin, R. Knox, M. Linzer, A. Neori and M. Olsson for technical assistance, and C. Eriksen, J. Barry, L. Fogiel, D. Genin, R. Hessler, T. Klinger, R. Knox, M. Mullin, L. Mullineaux, K. Smith, M. Tegner and W. Wakefield for discussions and comments on the manuscript. This research was supported by NSF grant OCE 85-10057, ONR contracts N00014-79-C-0472 and N00014-82-K-0147, and by a grant from The Scripps Industrial Association.

Received 14 February; accepted 7 May 1986.

1. Lonsdale, P., Normark, W. R. & Newman, W. A. *Bull. geol. Soc. Am.* **83**, 289-315 (1972).
2. Hogg, N. G. *J. Fluid Mech.* **58**, 517-537 (1973).
3. Roberts, D. G., Hogg, N. G., Bishop, D. G. & Flewelling, C. G. *Deep-Sea Res.* **21**, 175-184 (1974).
4. Huppert, H. E. *J. Fluid Mech.* **67**, 397-412 (1975).
5. Huppert, H. E. & Bryan, K. *Deep-Sea Res.* **23**, 655-679 (1976).
6. Owens, W. B. & Hogg, N. G. *Deep-Sea Res.* **27**, 1029-1045 (1980).
7. Gould, W. J., Hendry, R. & Huppert, H. E. *Deep-Sea Res.* **28**, 409-440 (1981).
8. Eriksen, C. *J. geophys. Res.* **87**, 525-538 (1982).
9. Roden, G. I. & Taft, B. A. *J. geophys. Res.* **90**, 839-856 (1985).
10. Pratt, R. M. in *Deep-Sea Photography* (ed. Hersey, J. B.) 145-158 (Johns Hopkins University Press, Baltimore, 1967).

11. Raymore, P. A. *Pacif. Sci.* **36**, 15-34 (1982).
12. Grigg, R. W. *Mar. Ecol.* **5**, 57-74 (1984).
13. Genin, A. & Boehlert, G. W. *J. mar. Res.* **43**, 907-924 (1985).
14. Genin, A., Boehlert, G. W. & Lonsdale, P. F. *Eos* **65**, 969 (1984).
15. Fukasawa, M. & Nagata, Y. *J. oceanogr. Soc. Japan* **34**, 41-49 (1978).
16. Sebens, K. P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5473-5477 (1984).
17. Pequegnat, W. E. *Ecology* **45**, 272-283 (1964).
18. Grigg, R. W. & Maragos, J. E. *Ecology* **55**, 387-395 (1974).
19. Reid, J. L. Jr *Intermediate Waters of the Pacific Ocean* (Johns Hopkins University Press, Baltimore, 1965).
20. Spiess, F. N. & Tyce, R. C. *Scripps Instrn. Oceanogr. Ref.* 73-74 (1973).
21. Sara, M. & Vacelet, T. *Traite de Zool.* **3**, 462-576 (1973).
22. Loya, Y. *Bull. mar. Sci.* **26**, 450-466 (1976).
23. Wunsch, C. *J. Fluid Mech.* **35**, 131-144 (1969).
24. Cacchione, D. & Wunsch, C. *J. Fluid Mech.* **66**, 223-239 (1974).

Monocular aniseikonia: a motion parallax analogue of the disparity-induced effect

Brian Rogers

Department of Experimental Psychology, University of Oxford,
South Parks Road, Oxford OX1 3UD, UK

Jan Koenderink

Department of Medical Physics, University of Utrecht,
The Netherlands

Mayhew and Longuet-Higgins have recently outlined a computational model of binocular depth perception¹ in which the small vertical disparities between the two eyes' views of a three-dimensional scene are used to determine the 'viewing parameters' of fixation distance (d) and the angle of asymmetric convergence of the eyes (g) (refs 2, 3). The d/g hypothesis, as it has been called⁴, correctly predicts that a fronto-parallel surface, viewed with a vertically magnifying lens over one eye, should appear to be rotated in depth about a vertical axis^{1,3-5}. We report here a comparable illusion for surfaces specified by monocular motion parallax information, which can be explained more simply by considering the differential invariants of the optic flow field. In addition, our observations suggest that the disparity-induced effect is not a 'whole field' phenomenon nor one limited to small magnification differences between the eyes^{1,4}.

The vertical magnification of one eye's image of a binocularly-viewed surface produces the impression of a surface rotated about a vertical axis through the fixation point. This effect was called the 'induced effect' by Ogle⁶ because he believed that horizontal disparities were 'induced' into the neural representations by a compensatory isotropic scaling mechanism acting to minimize the vertical size differences caused by eccentric fixation. Since horizontal disparities are affected by both the magnitude of any depth differences and the degree of eccentric fixation (g), the zero horizontal disparities in an induced effect stimulus have to be 'corrected' so that the relative distance and the slant of a surface can be perceived correctly. However, it appears that Ogle did not appreciate the more general significance that vertical disparities at other retinal locations apart from the vertical meridian provide a potential source of information about distance to the fixation point, as well as the angle of eccentric convergence. The mathematical proof for this has been independently provided by Gillam and Lawergren⁷ and Mayhew and Longuet-Higgins¹⁻³. In both cases, the induced effect is seen as a necessary consequence of a stereoscopic system which uses vertical disparities to determine the viewing system parameters.

As yet, there is little evidence that presence of vertical disparities in a stereogram actually gives a subjective impression of eccentric convergence but, as several authors have pointed out, this might be due to the presence of conflicting oculomotor information^{1,7}. However, the magnitude and the direction of the induced effect are both consistent with the d/g hypothesis³. Mayhew and Longuet-Higgins also argue the case for the induced effect being a global or 'whole field' phenomenon, since there could only be a single estimate of eccentric fixation angle and therefore only a single 'correction' would be applied to the horizontal disparities in the surrounding area^{1,4}. As additional evidence for their theory, they note that the magnitude of the induced effect does not increase when the vertical size difference is greater than about 4–6%^{1,6}. This would be predicted by their hypothesis since larger differences in size would imply impossibly large angles of asymmetric convergence^{1,7}. Both claims are examined in this paper.

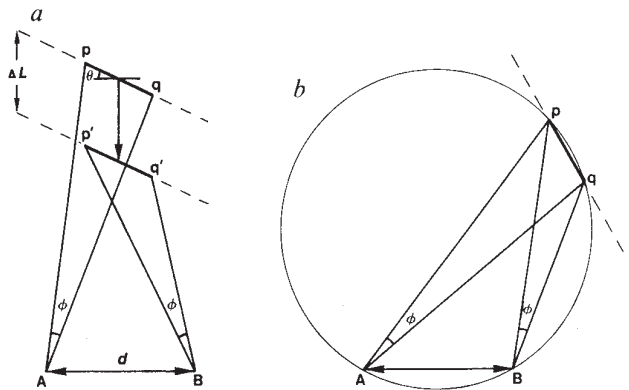


Fig. 1 In *a*, pq represents part of a surface slanting at an angle θ to the frontoparallel plane, which moves forward to $p'q'$ at the same time as a monocular observer moves from A to B . It can be shown that the horizontal angle subtended at the eye ϕ will remain constant when $\tan \theta = \Delta L/d$. In contrast, since the surface is closer at B than A , the vertical angle subtended will increase. In *b*, the horizontal angle subtended again remains constant and the vertical angle subtended increases when the monocular observer moves from A to B towards an eccentrically placed slanting surface pq . The two situations differ only by an eccentric rotation of the lines of sight.

The disparity-induced effect can be perceived when one eye's image of a binocularly-viewed scene is magnified vertically. The motion parallax analogue of this illusion—monocular aniseikonia—was created by continuously magnifying and minifying the image projected to a single viewing eye during side-to-side movements of the observer's head⁸. The instantaneous monocular images of the scene at the end points of the lateral head movement necessarily correspond to the two simultaneous binocular views of the same scene when a meridional magnifying lens is placed over one eye. For all the observations reported in this paper, the image transformations for both the parallax and disparity-induced effects were effected electronically, rather than by optical means. The images consisted of either a single random dot pattern, or a pair of random dot patterns viewed independently by the two eyes, each subtending a $20^\circ \times 20^\circ$ visual angle. The disparity-induced effect was produced by increasing the vertical gain slightly on one oscilloscope and decreasing it slightly on the other. The horizontal widths of the patterns remained identical. To produce the parallax-induced effect, the vertical gain of a single, monocularly-viewed oscilloscope was modulated according to the position of the observer's head. Thus the vertical size of the dot pattern was maximal at one end of travel and minimal at the opposite end of travel. Observers were asked to report the perceived orientation and shape of the random dot surfaces in the different experimental conditions.

With binocular viewing of the disparity-induced surface, our results replicate those of previous studies^{3,6,7}. Observers reported that the surface appeared to be slanting in depth with the right-hand side apparently closer than the left, when the right eye's image was vertically magnified, and vice versa for magnification of the left eye's image. Increasing the difference in vertical magnification between the two eyes increased the angle of perceived slant, as found previously.

With monocular viewing of the parallax-induced surface, a similar pattern of results was obtained. Observers reported that the surface appeared to be slanting in depth with the right-hand side closer than the left when the monocular image was progressively magnified with head movement to the right and vice versa. Again, the angle of perceived slant increased with an increase in the extent of vertical magnification/minification (for a constant head movement). In both the disparity-induced effect and its parallax analogue, the apparently slanting surfaces were

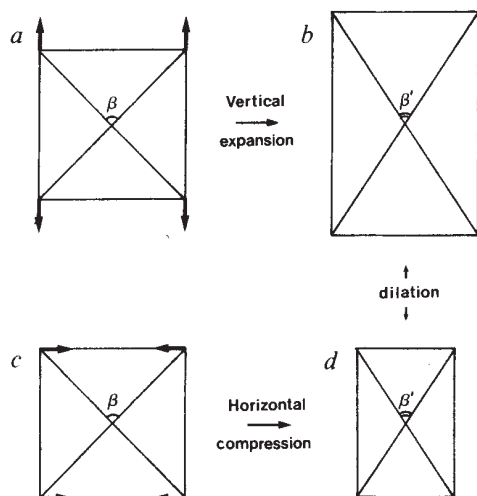


Fig. 2 The vertical expansion of the square pattern segment from *a* to *b* will produce the same amount of local and global deformation (as indicated by the change in the orientation difference between the diagonals $\beta \rightarrow \beta'$) as the horizontal compression of the same pattern segment from *c* to *d*. The transformed patterns, *b* and *d*, are related by a uniform or isotropic expansion.

always perceived as lying directly in front of the observer, rather than eccentrically to one side. In the case of the parallax-induced effect, the slanting surface was also perceived to approach and recede along a median path as the image expanded and contracted. This percept is entirely consistent with the geometry of the transformation, since the image of a slanting surface which approached an observer as he moved laterally towards the "closer" side, would indeed expand vertically, but remain of constant width (Fig. 1*a*).

The interpretation chosen by the visual system is, however, not the only one consistent with the image transformation. In fact, there are an infinite number of possible solutions (unlike the disparity-induced effect which has only one solution). For example, a surface positioned eccentrically and slanting with respect to the direction of gaze, would also generate an image whose vertical size increased and decreased with head movements whilst its horizontal size remained constant (Fig. 1*b*). This particular solution lies behind the explanation proposed by Mayhew and Longuet-Higgins to account for the disparity-induced effect¹. The fact that observers do not experience this perceptual outcome suggests that, for the parallax system, a vertical size change which accompanies a horizontal head movement is interpreted as a motion in depth rather than resulting from eccentric fixation.

Two additional sets of observations were made for the disparity- and parallax-induced effects. First, as mentioned above, the *d/g* hypothesis predicts that the induced effect should be a 'whole field' phenomenon, since it seems unlikely that the visual system could entertain different, and therefore contradictory, estimates of the angle of eccentric convergence at the same time. Empirical evidence for the 'whole field' characteristic of the effect comes from the observation that an embedded region which is magnified in one eye's view does not appear to be slanted with respect to the surround^{1,5,9}. However, the perception of opposite induced effects in neighbouring spatial regions is possible in both the classical induced effect and the parallax analogue reported here. For example, if the images of the left and right halves of the random dot pattern seen by the right eye are minified and magnified respectively (with respect to the images seen by the left eye), subjects report that the left half of the pattern appears to be slanting closer to the left and the right half slanting closer to the right. According to the *d/g* hypothesis this could only be possible if the vertical disparities in the left half field were interpreted as indicating asymmetric convergence

to the right, and in the right half field as indicating asymmetric convergence to the left. A comparable 'double' induced effect was also obtained for surfaces specified by motion parallax information.

The second piece of evidence cited in favour of the *d/g* hypothesis is that the induced effect does not increase with vertical magnifications greater than 4–6%^{3,6}. Indeed, this prediction was used by Frisby⁴ to discount Westheimer's finding that sensitivity to vertical disparities is much poorer than for horizontal disparities¹⁰. However, we have found that the apparent slant of induced effect surfaces does still increase monotonically right up to a 50% magnification difference, when induced effects specifying opposite slants are alternated every few seconds. At the 57-cm viewing distance used in the present experiments, the maximum possible magnification difference which could result from asymmetric convergence would be less than 7%.

Our proposed explanation of the parallax-induced effect is based on the use of differential invariants to describe the optic flow field^{11–13}. As Koenderink has shown, the slant of a surface is uniquely specified by the amount of deformation or shear in the flow field, if the extent of observer motion is known. Clearly, the vertical expansion of a monocular image, which underlies the parallax-induced effect, will produce the same degree of deformation in the flow field as a horizontal contraction of the same image (Fig. 2). Hence, the same surface slant is specified. What remains after the deformation component has been extracted is a simple divergence term which is positive in the first case and negative in the second. In the parallax-induced effect, the divergence term is clearly not ignored but instead is responsible for the apparent approach and recession of the slanted surface noted earlier.

Could the proposed explanation also account for the disparity-induced effect? Given that both the mathematical analyses of disparity and parallax transformations¹⁴ and their perceptual characteristics are so similar^{15,16}, it is tempting to speculate that the perceived surface slants seen in the disparity-induced effect are the result of having a visual system which uses the amount of deformation between the two binocular images as an indicator of surface slant¹⁷. If this were the case, then the visual system would be left to account for the isotropic size difference between the eyes (the divergence component). The human visual system may have evolved to use this as an indicator of eccentric fixation or it may simply be ignored. Hence, our explanation of the induced effect also allows for the angle of eccentric fixation to be recovered from the disparity field, but differs from the Mayhew and Longuet-Higgins interpretation in that this does not have to be done, even implicitly. According to the *d/g* hypothesis, vertical disparities are used to signal the angle of asymmetric convergence which is then used to scale horizontal disparities. The fact that vertical disparities do not appear to give the impression of asymmetric convergence, together with the other characteristics of the induced effects reported here, suggests that our proposed explanation may be more parsimonious.

This work was supported by the SERC (UK) and the Dutch Organisation for the Advancement of Pure Research (ZWO).

Received 27 January; accepted 13 April 1986.

1. Mayhew, J. E. W. & Longuet-Higgins, H. C. *Nature* **297**, 376–379 (1982).
2. Longuet-Higgins, H. C. *Perception* **11**, 377–386 (1982).
3. Mayhew, J. E. W. *Perception* **11**, 387–403 (1982).
4. Frisby, J. *Nature* **307**, 592–593 (1984).
5. Stenton, S. P., Frisby, J. P. & Mayhew, J. E. W. *Nature* **309**, 622–623 (1984).
6. Ogle, K. N. *Researches in Binocular Vision* (Saunders, Philadelphia, 1950).
7. Gillam, B. & Lawergren, B. *Percept. Psychophys.* **24**, 121–130 (1983).
8. Rogers, B. J. & Graham, M. E. *Perception* **8**, 125–134 (1979).
9. Mayhew, J. E. W. & Frisby, J. P. *Vision Res.* **22**, 1225–1228 (1982).
10. Westheimer, G. *Nature* **307**, 632–634 (1984).
11. Koenderink, J. J. & van Doorn, A. J. *Optica Acta* **22**, 773–791 (1975).
12. Koenderink, J. J. *Brain Mechanisms and Spatial Vision* (eds Ingle, D., Jeannerod, M. & Lee, D.) 31–58 (Nijhof, Amsterdam, 1985).
13. Koenderink, J. J. *Vision Res.* **26**, 161–179 (1986).
14. Longuet-Higgins, H. C. *Nature* **293**, 133–134 (1981).
15. Rogers, B. J. & Graham, M. E. *Vision Res.* **22**, 261–270 (1982).
16. Rogers, B. J. & Graham, M. E. *Science* **221**, 1409–1411 (1983).
17. Koenderink, J. J. & van Doorn, A. J. *Biol. Cybernet.* **21**, 29–35 (1976).

New HLA DNA polymorphisms associated with autoimmune diseases

H. Festenstein*, J. Awad*, G. A. Hitman†, S. Cutbush*, A. V. Groves*, P. Cassell‡, W. Ollier* & J. A. Sachs*‡

* Department of Immunology, † Medical Unit and ‡ Bone and Joint Research Unit, London Hospital Medical College, Turner Street, London E1 2AD, UK

Certain class II determinants of the human histocompatibility locus antigens (HLA) have been implicated in the aetiology of several autoimmune diseases, including rheumatoid arthritis (RA) and insulin-dependent diabetes mellitus (IDDM). HLA-Dw4 was the first HLA determinant found to be significantly increased in RA patients compared with controls¹, while Dw4 and Dw3 were found to be significantly increased in IDDM patients^{2,3}. When the HLA-DR system was defined, RA patients were found to have an increased frequency of DR4 and IDDM patients an increased incidence of both DR4 and DR3 (ref. 4) compared with controls. As the HLA-Dw specificities are narrower than the serologically defined DR specificities, it was of specific interest to the present study that Dw4, Dw10, Dw13, Dw14, Dw15 and DKT2 are included in DR4 (ref. 5). We describe here new restriction fragment length polymorphisms (RFLPs) and, together with the newly described serologically defined DQ specificity TA10 (ref. 6), test their prevalence and associations in controls and diseased patients. We find that the newly characterized DNA bands are present at a much higher frequency in RA and IDDM patients than in controls. These findings may lead to a greater understanding of the pathogenesis of such diseases.

The HLA class II region is considered to consist of five α -chain genes and eight β -chain genes; these are encoded in three subregions, HLA-DP, -DQ and -DR. The DR subregion contains one α -chain gene and three or four polymorphic β -chain genes, while the DP and DQ subregions each have two α - and two β -chain genes, all of which are polymorphic. The class II molecules encoded by these genes are transmembrane dimers consisting of an α - and a β -chain. Certain alleles of the HLA-DR series are in strong positive linkage disequilibrium with alleles of the DQ series (reviewed in ref. 7). The genetic basis of the Dw determinants, determined by homozygous typing cells (HTCs) in mixed lymphocyte culture, remains controversial. It is our view that these determinants are encoded in the HLA-DR subregion but are not identical to the specificities of the DR segregant series^{8,9}.

We first investigated the DR4 RFLP, using as reference DNA panels of HTCs covering all the Dw specificities. The DNA was digested with *TaqI*, *PvuII*, *PstI* and *EcoRI*, electrophoresed on

agarose gels, and the gels were then blotted onto nitrocellulose filters and hybridized with DR β , DQ β and DQ α probes. Out of 11 DR4⁺ HTCs tested, 3 were Dw4, 1 was Dw10, 2 each were Dw13, Dw14 and Dw15 and 1 was DKT2 (Table 1).

No differences were seen in the Southern blot patterns obtained after hybridization with the DR β probe (not shown). However, after *TaqI* digestion followed by hybridization with the DQ α probe, which detects DX and DQ genes^{10,11}, we observed clear differences in the Southern blot bands. This probably represents the DX α polymorphism first described by Hui *et al.*¹² and confirmed by others¹³. *TaqI* digestion showed the DX α polymorphism to consist of 2.1-kilobase (kb) ('upper' = U) and 1.9-kb ('lower' = L) fragments; U and L correspond to the 2.2- and 2.1-kb bands described by Nepom *et al.*¹⁴. Two consanguineous DR4/Dw4 HTCs tested show different patterns; MCF contains the L fragment while BM14 contains the U fragment. All the other HTCs have one or other of these fragments; the two Dw13 HTCs also have different DX α RFLPs (Fig. 1). In addition, BM14 and MCF give different DQ α RFLPs using *PvuII* (8 and 7.5 kb respectively, not shown); these probably also represent the DX α polymorphism, and we have now used this to investigate HLA associations in IDDM and RA. Our results suggest that DQ differences do not correlate with Dw definitions. Figures 2 and 3 show Southern blot patterns obtained using the DQ β probe and *EcoRI*, *PstI*, *PvuII* and *TaqI* to digest the DNA. Two patterns, which we have called Ω and ϕ (see Table 1), were seen among the HTCs; when they are compared using BM14 (Ω) and MCF (ϕ) (Fig. 2), both of which are DR4/Dw4, the bands which distinguish most clearly between the two patterns are of 12 kb (Ω) and 4.4 kb (ϕ) using *PstI*, and 19 kb (Ω) and 13 kb (ϕ) using *EcoRI*, while with *PvuII* a 2.6-kb band is absent in Ω and present in ϕ . The patterns obtained on *TaqI* digestion (Fig. 3) are more complex, consisting of bands of 2.0, 2.3 and 2.5 kb for Ω and of 1.8, 1.95, 2.2, 2.4 and 2.6 kb for ϕ . The two DR4/Dw13 HTCs show different DQ β RFLPs, as do the two DR4/Dw4 HTCs (Fig. 3). These results support our earlier findings that the DQ types detected using alloantisera do not materially affect HLA-Dw assignment and that the same Dw specificity can occur in cells that differ for HLA-DQ (ref. 15).

We next performed serological and monoclonal antibody studies of the DR4 cells using the FLEGG alloantiserum¹⁶ and the A10-83 monoclonal antibody (Maeda, 9w790) in a standard microcytotoxicity test. These reagents define a specificity designated TA10 (ref. 6) which splits DQw3 into DQw3.1 and DQw3.2 (see ref. 14) and which is associated with DR5-DQw3 and a subset of DR4. One of the two DR4/Dw4 HTCs was TA10-positive (MCF, ϕ) and the other (BM14, Ω) was TA10-negative (Table 1). The two DR4/Dw13 HTCs also had disparate TA10 typings. Table 2 shows the frequency of TA10 positivity in the DR4⁺, DR5⁻ controls. Positivity for the TA10 subtype of DQw3 coincides with the DQ β ϕ pattern in the small number of DR4⁺

Table 1 DR4⁺ homozygous typing used in the present study

Name (source)	HLA							RFLP		
	A	B	Cw	Dw	DR	DQw	TA10	DX	DQB Ω or ϕ	DQB β T6
MCF (Local)	2/-	62/-	3/-	4/-	4/-	3/-	+	L	ϕ	+
BM14 (Ferrara, Bergamo)	3/-	7/-	?	4/-	4/-	3/-	-	U	Ω	+
I9W4(ER) (Dupont, New York)	2/-	44/-	5/-	4/-	4/-	3/-	NT	NT	ϕ	NT
I10W10(FS) (Dupont, New York)	26/-	38/-	?	10/-	4/-	3/-	NT	NT	Ω	+
FLE (Local)	2/-	63/-	7/-	13/-	4/-	3/-	-	U	Ω	+
JHA (Local)	31/-	51/-	-/-	13/-	4/-	3/-	+	L	ϕ	+
MT (Bashir, Sydney)	31/-	60/-	3/-	14/-	4/-	3/-	NT	U	Ω	+
KY (Bashir, Sydney)	31/-	60/-	3/-	14/-	4/-	3/-	NT	U	Ω	+
EbWa (Kashiwagi, Kanagawa)	24/-	54/-	1/-	15/-	4/-	Blank	NT	L	Ω	+
KT9 (Kashiwagi, Kanagawa)	11/24	40/-	1/3	15/-	4/-	Blank	NT	NT	Ω	+
KT13 (Kashiwagi)	24/31	7/51	-/-	DKT2	4/-	?	NT	NT	Ω	+

NT, not tested.

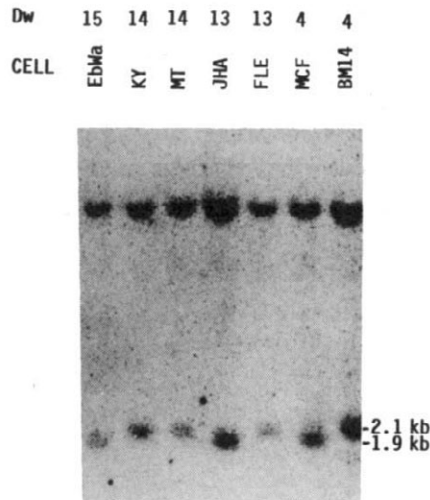


Fig. 1 DNA of homozygous DR4 cells digested with *TaqI* and probed with *DQα*. The figure shows the DXα 'U' band at 2.1 kb and the 'L' band at 1.9 kb. The cell lines MCF, JHA and EbWa have the L band, whereas BM14, FLE, MT and KY have the U band.

Methods. Genomic DNA was either extracted from Epstein-Barr virus-transformed cell lines derived from the HTCs shown in Table 1, or from peripheral blood lymphocytes from patients and controls. DNA (10 µg) was digested for 16 h with *EcoRI*, *PvuII* or *PstI* at 37 °C, or with *TaqI* at 65 °C (enzymes from Boehringer Mannheim). The resulting DNA fragments were separated on 0.5% agarose gels by electrophoresis at 30 V for 18–20 h and then transferred³⁰ to nitrocellulose filters (Schleicher & Schuell). Fragments of *HindIII*-digested λ DNA were used as molecular size markers. The DNA probes used were labelled with [α -³²P]dATP and [α -³²P]dTTP (NEN) by nick-translation³¹. Before hybridization, the nitrocellulose filters were incubated for 2 h at 65 °C in 0.45 M NaCl/45 mM trisodium citrate (pH 7.0) and 10× Denhardt's reagent. The filters were then hybridized with the ³²P-labelled heat-denatured probe in 0.45 M NaCl/45 mM trisodium citrate (pH 7.0)/10× Denhardt's reagent/10% sodium dextran sulphate and 500 µg denatured salmon sperm DNA per ml of hybridization mix. After 16 h at 65 °C the filters were washed three times at 65 °C for 15 min each with 0.45 M NaCl/45 mM trisodium citrate (pH 7.0)/0.1% sodium dextran sulphate, three times at 65 °C for 15 min each with 0.15 M NaCl/15 mM trisodium citrate (pH 7.0)/0.1% sodium dextran sulphate, and three times at 65 °C for 15 min each with 50 mM NaCl/5 mM trisodium citrate (pH 7.0)/0.1% sodium dextran sulphate. The labelled fragments were then detected by exposing Kodak XAR-5 film with intensifying screens to the radioactive filters at -70 °C for 2–3 days.

HTCs. This was confirmed by Nepom *et al.*¹⁷ using a larger panel and may correspond to the DQR4 (Ω) and DQR5 (φ) patterns described by Cohen-Haguenauer *et al.*¹⁸. Although the TA10/*DQβφ* association is not absolute in heterozygous individuals, as TA10 is included within *DQβφ*, it may define a *DQβφ* subtype (Table 3). TA10 positivity is also associated with B44 in our DR4⁺ panel.

As all these new polymorphisms are associated with each other and with specific HLA determinants, we were able to construct three series of associated alleles, or preferential allelic associations (PAAs); one was based on the *B8/DR3* haplotype, one on the *DR4/B44* haplotype and one on the *DR4/Bw62* haplotype, which includes alleles not only of the newly described RFLPs (see below) but also of TA10.

Disease associations were thus investigated in two ways. First, the incidence of the newly described polymorphisms was compared directly in patients and control subjects—DXαU and L; *DQβφ* and Ω; and *DQβT6* and TA10 positivity and negativity. *DQβT6* was defined by a 6-kb fragment obtained with the *DQβ* probe following *TaqI* digestion and was present in all the DR4 HTCs (Fig. 3) but was absent in some of the DR4⁺ patients and controls.

In diabetics who type as DR4, the positive associations between diabetes and certain RFLPs compared with those of control subjects are as follows: DXαU, 92% versus 46%; *DQβφ*, 81%, 46%; *DQβT6*, 81%, 24%; and Bw62, 39%,

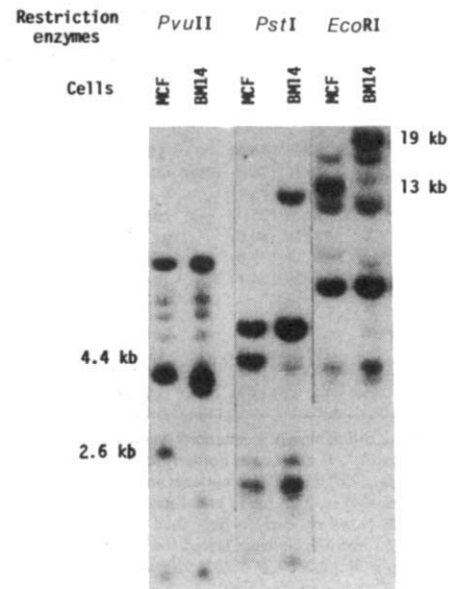


Fig. 2 DNA from HTCs BM14 and MCF (DR4, Dw4) hybridized with a *DQβ* probe. The figure shows the distinguishing features between the *DQβΩ* pattern (BM14) and the *DQβφ* pattern (MCF). *EcoRI* digestion gives a 19-kb (Ω) and a 13-kb (φ) band, *PstI* a 12-kb (Ω) and a 4.4-kb (φ) band, and with *PvuII* a 2.6-kb band is present in the φ pattern and absent in the Ω pattern. When DNA from DR4 individuals is digested with *PstI*, the 12.5- and 4.4-kb bands are used to distinguish Ω from φ. However, other DR types also give one or other of these bands; for example, DR1 has a 4.4-kb band, DR2 and DR6 give a 12.5-kb band. It is therefore possible to designate the DR4 to Ω or φ by eliminating a band associated with the other haplotype. For example, a DR1, DR4 patient displayed both the 12.5- and 4.4-kb bands. By eliminating the 4.4-kb band associated with the DR1 it is possible to assign the DR4 to the Ω type. On this basis all individuals with DR1, DR2 or DR6 showing both bands were classified as Ω or φ. There was one DR4, DR7 patient and two controls not classified due to clear polymorphism in the DR7. Also amongst the controls was one DR4, DR11 which was not classified due to insufficient data on this group, and a DR3, DR4 due to possible polymorphism here also.

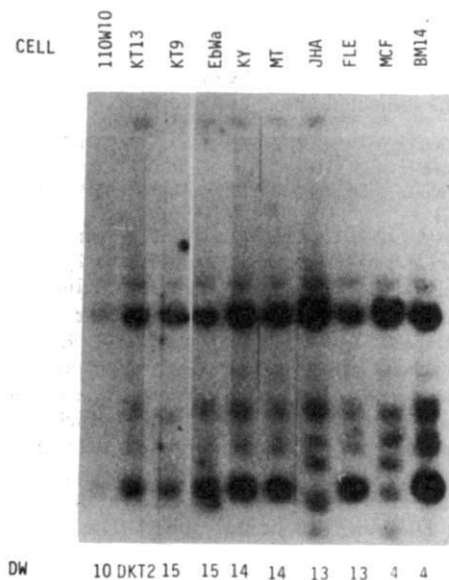


Fig. 3 DNA from HTCs BM14 and MCF (DR4, Dw4) hybridized with a *DQβ* probe. The distinguishing features between *DQβΩ* and φ patterns with *TaqI* are 2.0-, 2.3- and 2.5-kb bands with Ω (BM14) and 1.8-, 1.95-, 2.2-, 2.4- and 2.6-kb bands with the φ (MCF) pattern. Note that the 6-kb band is present in all the DR4 cell lines. All cells give the Ω pattern except for MCF and JHA.

Table 2 Frequencies in DR4⁺ controls, RA and IDDM patients, of DX α and DQ β RFLPs, TA10 and HLA-Bw62 and HLA-B44

RFLP	Controls (%)	RA patients (%)	IDDM patients (%)
DX α U	13/28 (46.4)	15/29 (51.7)	47/51 (92.2)
DX α L	21/28 (75.0)	21/29 (72.4)	24/51 (47.1)
DQ β Ω *	12/26 (46.1)		29/36 (80.5)
DQ β Ω †	6/24 (25.0)	16/30 (53.0)§	
DQ β ϕ *	21/26 (80.8)		12/27# (44.0)
DQ β Ω †	19/24 (79.0)	18/30 (60.0)	
DQ β T6	7/29 (24.1)	18/34§ (52.9)	29/36 (80.5)
Serological specificities			
TA10‡	10/24 (41.7)	13/27 (48.1)	4/15 (26.7)
Bw62	5/30 (16.7)	10/32 (31.2)	18/46¶ (39.1)
B44	13/30 (43.3)	13/32 (40.6)	8/46¶ (17.4)

Significance of differences in frequencies was tested by Haldane's modification²⁸ of Woolf's method²⁹. *P* values are not corrected for the number of tests.

* Assignments of DQ β in IDDM patients and control subjects were based solely on *Taq*I RFLPs, that is, 2 kb for Ω and 1.9 kb for ϕ .

† Assignments based on *Pst*I, as described in Fig. 2 legend.

‡ DR4⁺, DR5⁺ individuals were tested for TA10.

|| *P* < 0.0005; # *P* < 0.01; ¶ *P* < 0.05; § *P* = 0.06.

17%; while negative ones are with DX α L, 47% versus 75%; DQ β ϕ , 44%, 81%; B44, 17%, 43%; and TA10, 27%, 42%. In DR3⁺ diabetics there is also a strong association with DX α U (Table 2).

In rheumatoid arthritis patients, there were no significant differences in the DX α U and DX α L frequencies in DR4⁺ patients and control subjects, although associations were found with DQ β Ω (53% compared with 25%), DQ β T6 (53%, 24%) and Bw62 (31%, 17%) (Table 2).

We next investigated HLA haplotypes that might be potential disease markers by testing the associations between the newly described RFLPs, the serologically detectable DQ-like determinant TA10 and established antigens of the DR, DQ, Dw and HLA-B segregant series in the IDDM and RA patients. Although additional DR4⁺ IDDM patients and controls were used in this analysis, the total control group does not constitute a random panel, having a preponderance of DR3⁺ and DR4⁺ individuals. The significant associations (Table 3) are DX α U with DQ β Ω , DX α L with DQ β ϕ , TA10 with DQ β ϕ and TA10 with B44; DX α U with DR3; DQ β T6 with DR4 and DQ β T6 with DR1. However, the significance of these associations varied in the three groups tested (control subjects, RA and IDDM patients). The constructs of the three PAAs are based on the significant associations indicated above and shown in Table 2; they are:

PAA 1. DX α U ······ DR3 · Dw3 · B8
 PAA 2. DX α U · DQ β Ω · TA10⁺ · DQw3 · DR4 · Dw4 · Bw62
 PAA 3. DX α L · DQ β ϕ · TA10⁺ · DQw3 · DR4 · Dw4 · B44

The DX α U/DQ β Ω or DX α L/DQ β ϕ association was present in all the HTCs except EbWa, Dw15 (Table 1).

Our main aim was to try to identify specific molecules or haplotypes likely to be directly involved in the pathogenesis of RA and IDDM. The striking association (>92%) of DX α U with IDDM may be attributable to the high prevalence of the combinations of PAA 1 and PAA 2 which have previously been marked by DR3 and DR4 antigens, both of which are associated with DX α U and are present at a high frequency in IDDM³². It is already known that individuals heterozygous for DR3 and DR4 run a higher risk of developing IDDM than do those who are either DR3 or DR4 homozygous¹⁹. The DX α U allele in itself may thus be a risk factor for IDDM. Our studies also indicate that the Bw62/DR4 haplotype is important in both IDDM and RA. The RA patients used in our study did not show as marked an increase in the DQ β Ω component that is part of PAA 2 as did the controls, even though increased frequencies of the Bw62/DR4 haplotype have been reported in RA patients with severe disease^{20,21}. Our RA patients, although fulfilling the clinical criteria for RA, were not selected for severe

Table 3 Associations between RFLPs and serological specificities in controls, RA and IDDM patients

RFLP versus RFLP	<i>n</i>	++	+-	-+	--	<i>P</i>
DQ α versus DQ β						
DX α U versus DQ β Ω * in:						
Controls	23	6	7	0	10	0.017
RA patients	28	14	5	1	8	0.003
IDDM patients	32	19	8	2	3	0.21
DX α L versus DQ β ϕ * in:						
Controls	23	18	2	0	3	0.0056
RA patients	30	17	6	0	7	0.0008
RFLP versus class II antigens						
DQ β						
DQ β ϕ versus TA10* in:						
Controls	16	8	5	0	3	0.10
RA patients	21	9	2	2	8	0.007
DQ β T6 versus DR4† in:						
Controls	44	8	4	21	11	0.62
RA patients	62	23	10	19	10	0.47
IDDM patients	46	28	1	7	10	<0.00005
DQ β T6 versus DR1† in:						
Controls	36	5	3	0	28	0.00015
RA patients	62	19	14	0	29	1.9 × 10 ⁻⁷
IDDM patients	46	4	25	1	16	0.38
DQ α U						
DX α U versus DR3† in:						
Controls	48	17	9	1	21	<0.00005
IDDM patients	68	41	21	1	5	0.03
RA patients	30	1	18	1	10	0.61
Class I versus class II						
B44 versus TA10 in						
DR4 controls	24	8	3	3	10	0.02

P values were computed using Fisher's exact test.

* DR4⁺ individuals; † all individuals.

disease, and greater increases in DQ β Ω and DX α U might thus be expected in patients with severe disease. There was no difference in the frequency of TA10 in the RA patients compared with the controls, but its frequency was reduced in IDDM patients, as found by Tait *et al.*²².

We have also found an increase in the frequency of DQ β Ω in IDDM patients. However, Cohen-Haguenauer *et al.*¹⁸ reported that, compared with DR-matched controls, IDDM patients showed a decrease in a *Taq*I, DQ β , 2.1-kb fragment, which may correspond to our DQ β Ω pattern in PAA 2. This difference may be due to heterogeneity between the populations studied.

The increase in the frequency of the DQ β T6 RFLP in the RA and IDDM patients seems to be a novel finding, whereas the increases in DX α U and DQ β Ω in the IDDM patients represent extensions of known susceptibility axes. In IDDM patients the DQ β T6 fragment is found at a very high frequency, together with DR4 and DQ β Ω (Table 3); these associations are not apparent in the controls and hence it is not certain whether they represent true associations or combined increases of separate elements in the IDDM patients. The DQ β T6 RFLP does not show preferential association with Bw62 in the DR4⁺ IDDM patients (7/8 of the Bw62⁺ patients and 15/19 of the Bw62⁻ patients have the DQ β T6 fragment). As the frequency of the DQ β T6 RFLP is also increased in the RA patients compared with the controls, it could be a new common susceptibility marker in these disorders. Our data also indicate a highly significant association between DQ β T6 and DR1, especially in RA patients (Table 3); this may have some bearing on the reported increased frequency of DR1 in British²³, Indian²⁴ and Jewish²⁵ RA patients. Interestingly 'MC1', a specificity common to DR1 and DR4, has been described²⁶ and a monoclonal antibody recognizing this specificity has been produced which reacts with cells from more than 90% of RA patients (R. Winchester, personal communication).

Although it is not known whether the DX α U and L, DQ β Ω and ϕ DQ β T6 fragments encode functional HLA class II molecules, we can postulate that these RFLPs may encode class II molecules associated with the pathogenesis of IDDM and RA. Alternatively, they may be in linkage disequilibrium with

some other, more relevant, genes in the PAAs concerned. We consider that there is not a single gene and/or gene product of each of the PAAs involved, but that several different genes or molecules, possessing complementary functions with hierarchical as well as additive effects, are implicated. Hybrid DQ molecules²⁷, *trans* combinations of DQ α chains from one PAA and DQ β chains from another PAA, may also be relevant, particularly with regard to IDDM susceptibility. These findings provide a basis for investigating the putative HLA class II molecules potentially encoded by genes associated with the DQ RFLPs described here.

We thank Dr P. Petersen (DNA probes), Drs G. B. Ferrara, B. Dupont, H. Bashir and N. Kashiwagi (homozygous typing cells) and Dr H. Maeda (monoclonal antibody) for reagents used in this study. Patient material was provided by Professor H. Currey and Dr J. P. Monson of the London Hospital. Financial support was provided by the Arthritis and Rheumatism Council (S.C., W.O.), the Multiple Sclerosis Society (J.A., A.V.G.), MRC (H.F.) and the Wellcome Trust (J.A.S., P.C.). We thank Joanna Lee for help in the preparation of this manuscript.

Received 7 February; accepted 26 March 1986.

1. Stastny, P. *Tissue Antigens* 4, 571-579 (1974).
2. Thomsen, M. *et al. Transplant Rev.* 22, 120-147 (1985).

3. Sachs, J. A., Cudworth, A. G., Jaraquemada, D., Gorsuch, A. N. & Festenstein, H. *Diabetologia* 18, 41-43 (1980).
4. Batchelor, J. R. & Morris, P. J. in *Histocompatibility Testing* 1977 (eds Bodmer, W. F. *et al.*) 205-258 (Munksgaard, Copenhagen, 1977).
5. Jaraquemada, D. *et al.* in *Histocompatibility Testing* 1984 (eds Albert, E. D. *et al.*) 270-273 (Springer, Berlin, 1984).
6. Maeda, H. *Tissue Antigens* 23, 163-170 (1984).
7. Giles, R. C. & Capra, J. D. *Tissue Antigens* 25, 57-68 (1985).
8. Jaraquemada, D., Okoye, R. C., Ollier, W., Awad, J. & Festenstein, H. in *Histocompatibility Testing* 1984 (eds Albert, E. D. *et al.*) 472-473 (Springer, Berlin, 1984).
9. Sachs, J. A., Jaraquemada, J. & Festenstein, H. *Tissue Antigens* 17, 43-56 (1981).
10. Auffray, C., Kuo, J., Demars, R. & Strominger, J. L. *Nature* 304, 174-177 (1983).
11. Trowsdale, J. *et al. Proc. natn. Acad. Sci. U.S.A.* 80, 1972-1976 (1983).
12. Hui, K. M. *et al.* in *Histocompatibility Testing* 1984 (eds Albert, E. D. *et al.*) 590-594 (Springer, Berlin, 1984).
13. Spielman, R. S., Lee, J., Bodmer, W. F., Bodmer, J. G. & Trowsdale, J. *Proc. natn. Acad. Sci. U.S.A.* 81, 3461-3465 (1984).
14. Nepom, B. S. *et al. J. exp. Med.* (in the press).
15. Jaraquemada, D. *et al. Hum. Immun.* (in the press).
16. Navarrete, C. thesis, Univ. London (1985).
17. Kim, S. J. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
18. Cohen-Haguenauer, O. *et al. Proc. natn. Acad. Sci. U.S.A.* 82, 3335-3339 (1985).
19. Sveigaard, A., Platz, P. & Ryder, L. P. *Immun. Rev.* 70, 193-218 (1983).
20. Ollier, W. *et al. Tissue Antigens* 24, 279-291 (1984).
21. Jaraquemada, D. *et al. Ann. rheum. Dis.* (in the press).
22. Tait, B. D. *et al. Tissue Antigens* 24, 228-233 (1984).
23. Grennan, D. M., Sanders, P. A., Dyer, P. A. & Harris, R. *Ann. Rheum.* (in the press).
24. Nichol, F. E. & Woodrow, J. C. *Lancet* i, 220-221 (1981).
25. Schiff, B., Mizrahi, Y., Orgad, S., Yaron, J. & Gazit, E. *Ann. rheum. Dis.* 41, 403 (1982).
26. Duquesnoy, R. J., Marrari, M., Hackbarth, S. & Zeevi, A. *Hum. Immun.* 10, 165-176 (1984).
27. Charron, D. J., Lotteau, V. & Turmel, P. *Nature* 312, 157-159 (1984).
28. Haldane, J. B. S. *Ann. hum. Genet.* 20, 309-311 (1955).
29. Woolf, B. *Ann. hum. Genet.* 19, 251-253 (1955).
30. Southern, E. M. *J. molec. Biol.* 98, 503-517 (1975).
31. Rigby, P. W. J., Diekmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* 113, 237-251 (1977).
32. Hitman, G. A. *et al. Immunogenetics* 23, 47-51 (1986).

Polymorphism of human Ia antigens: gene conversion between two DR β loci results in a new HLA-D/DR specificity

Jack Gorski & Bernard Mach

Department of Microbiology, University of Geneva Medical School, 1211 Geneva 4, Switzerland

The polymorphic HLA-DR β -chains are encoded within the human major histocompatibility complex (MHC) by multiple loci resulting from gene duplications. Certain DR haplotypes can be grouped into families based on shared structural factors. We have studied the molecular basis of HLA-DR polymorphism within such a group which includes the haplotypes DR3, DR5 and DRw6. Molecular mapping of the DR β -chain region allows true allelic comparisons of the two expressed DR β -chain loci, DR β I and DR β III. At the more polymorphic locus, DR β I, the allelic differences are clustered and may result from gene conversion events over very short distances. The gene encoding the HLA-DR3/Dw3 specificity has been generated by a gene conversion involving the DR β I and the DR β III loci of the HLA-DRw6/Dw18 haplotype, as recipient and donor gene, respectively. Based on which allele is found at DR β III, the less polymorphic locus, two groups of haplotypes can be defined: DRw52a and DRw52b. The generation of HLA-DR polymorphism within the DRw52 supertypic group can thus be accounted for by a succession of gene duplication, divergence and gene conversion.

The polymorphism of class II molecules is responsible for the restriction of T-cell activation by antigen-presenting cells. The polymorphic regions of the class II molecules are therefore implicated in interactions with both the T-cell receptor and with antigen. Since some alleles are more efficient at presenting certain antigens than others, class II genes are also referred to as immune response genes (reviewed in refs 1 and 2). The class II products of the human MHC (the HLA complex) include the HLA-DR, DQ and DP antigens³. HLA-DR molecules are the predominant class II products on the cell surface, and the polymorphism is associated with the β -chain. The polymorphic regions of the DR β -chains are located predominantly at the N-terminal of the mature polypeptide chain⁴, corresponding to the first structural domain. The DR subregion contains multiple highly homologous β -chain loci⁵⁻⁷ but only one α -chain

locus⁸⁻¹⁰. In the DR3 and DRw6 haplotypes, the three β -chain loci have been linked on a molecular map¹¹. The DR β II locus is a pseudogene without a first domain exon and the two expressed loci, β I and β III, are the result of a recent duplication¹¹.

The serologically identified DR specificities can be grouped on the basis of their reactivity with supertypic sera¹². The haplotypes within such a supertypic group have similar Southern blot patterns^{13,14} and, within the supertypic group DRw52, restriction maps of genomic clones are very similar¹¹. This implies a common evolutionary origin of the haplotypes within a supertypic group. Based on the structural similarity of the HLA-DR β -chain region of these haplotypes, individual DR β -chain genes can now be assigned to specific loci and true allelic comparisons made.

The nucleotide sequences of the polymorphic first domains of the two active loci (β I and β III) of the DR3, DRw6a and DRw6b haplotypes were determined. The sequence of a DR5 β I locus¹⁵ is added to the comparison. The allelic sequences for each locus (Fig. 1a and b) are presented in relation to a consensus sequence. It is evident that locus DR β I (a) is more polymorphic than DR β III (b) as all the DR β I sequences are different. Within the locus DR β I there is a clearcut clustering of the nucleotide differences around position 185-205 (the 'hypervariable region'). This is in contrast with DR β -chain sequence comparisons made in a nonallelic manner and across supertypic groups, where a different pattern of variation is observed. This can be seen when the consensus sequences of the two loci are compared (Fig. 1c). An additional region of variability is seen from base pair (bp) 15 to 22.

The nature of the polymorphic differences at the β I locus was explored in detail. It was observed that the sequence of the DR β I locus of HLA-DR3 can be generated by exchanging two short segments of DNA from the β I and the β III loci of the haplotype DRw6a (Fig. 2a). Given the phenomenon of positive interference for multiple double recombination events, it is more likely that a single gene conversion with interrupted repair¹⁶ accounts for the structure of the DR3 β I locus. A schema of the conversion event is shown in Fig. 2b, where the DRw6a β III locus acts as donor and the DRw6a β I locus as recipient, resulting in the sequence of the DR3 gene. As expected with such a mechanism, the β III locus of HLA-DR3 is identical to that of HLA-DRw6a.

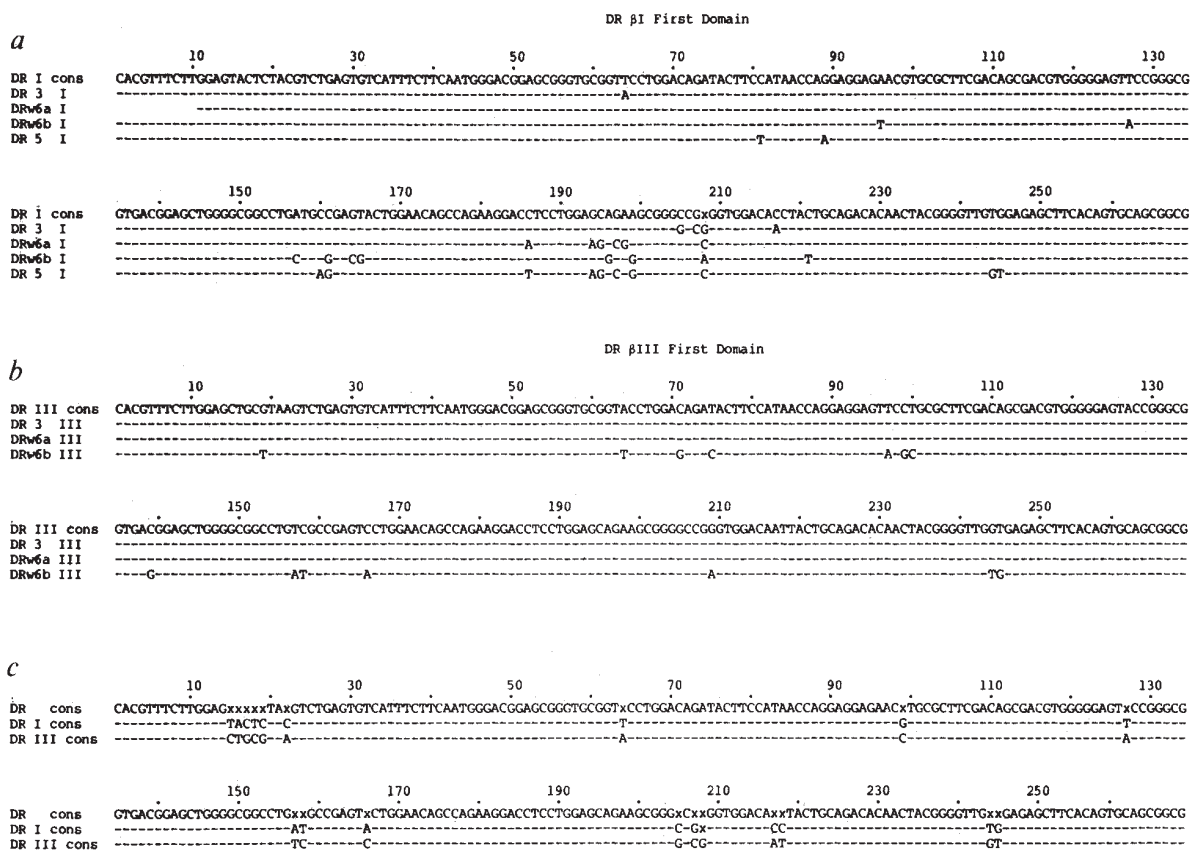


Fig. 1 Sequence of the DR β -chain first domain of the DR3 haplotype, two DRw6 haplotypes and a published DR5 haplotype¹⁵. **a**, β I locus; **b**, β III locus. The consensus sequence (cons) is the base found at that position in at least two of the sequences. **c**, Comparison of locus I and locus III sequence differences by aligning the consensus sequences for each locus and deriving a DR β consensus. Regions where β I and β III differ from each other and which may influence locus-specific conformation of the protein product are easily visualized. **Methods.** Sequences were obtained from subclones of the first domain exon of previously described cosmid and phage genomic clones^{6,11}, by both the chemical²³ and dideoxy²⁴ methods, as well as from cDNA clones⁵. DR3 sequences are from a cosmid library of the consanguineous homozygous typing cell line AVL; DRw6a sequences are derived from a cosmid library and cDNA library of the consanguineous homozygous typing cell line HHK; DRw6b sequences are from a Charon 30 phage library of a DR4, DRw6 line. The DRw6 genes were identified by Southern blot comparisons with DRw6 and DR4 genomic DNAs (see ref. 25 for an example). The cDNA-derived sequence of DRw6b β III⁵ as well as a phage-derived sequence²⁵ have been published previously.

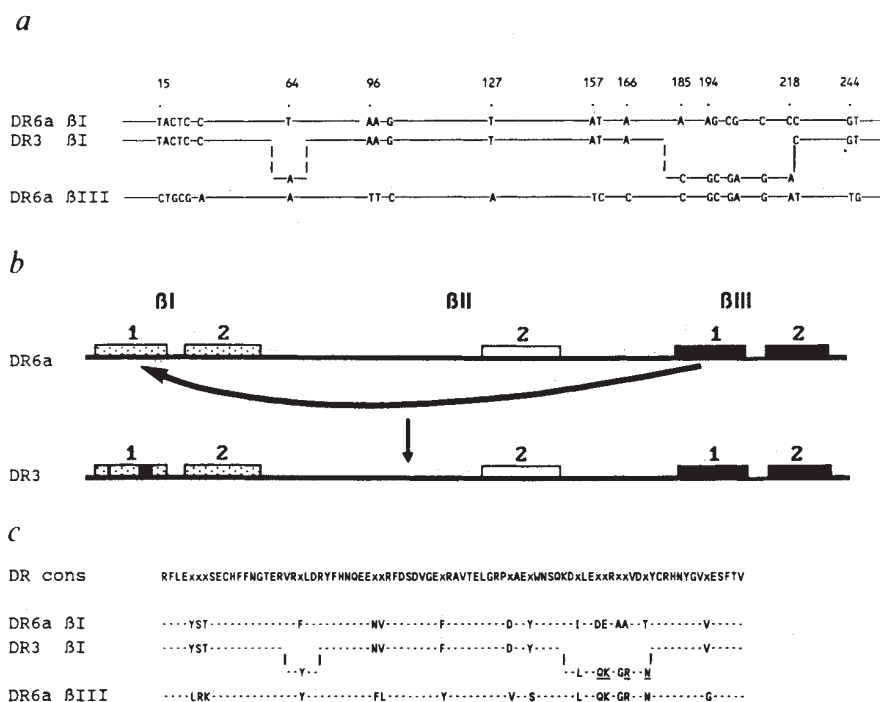


Fig. 2 Intrachromosomal gene conversion between two DRw6 β -chain loci genes. **a**, The first-domain sequences of the two DRw6 loci are aligned with the sequence of the DR3 β I locus. A solid line indicates identical sequences. Relevant nucleotide positions are shown to facilitate comparison with the sequence in Fig. 1. Because of the uncertainty in the limits of the conversion event, dotted lines are used to indicate cross-over points. **b**, Schematic representation of genes participating in the conversion event; **c**, amino-acid sequence of the first domains of the DR3 β I and DRw6a β I and β III chains showing the block of amino acids transferred by the conversion event. Non-conservative amino-acid changes in the DR3 β I chain are underlined. The consensus approach is used to aid visualization of amino-acid differences between the β I and β III loci. The block of amino acids transferred must interact with these differences in the generation of the novel epitope.

The replacement of this short segment of DNA of the *DRw6* βI locus (Fig. 2) has given rise to a gene encoding a different serological and T-cell specificity. The B-cell lines AVL (*DR3*, 3) and HHK (*DRw6*, w6), from which these genes have been cloned, have distinct serological reactivities, and alloreactive T-cell proliferation assays distinguish these lines as Dw3 and Dw18, respectively¹⁷. The transfer of the block of amino acids from the DR βIII -chain to the DR βI -chain (Fig. 2c) results in the loss of two negatively charged amino acids and a gain of two positively charged residues. The introduction of this region into the context of the βI locus must result in a novel conformation which gives rise to the DR3/Dw3 specificity. *DR3* is a common haplotype which is associated with an increased risk for certain diseases.

Upon further analysis of the highly polymorphic region of the *DRBI* locus, an example of a possible interchromosomal

	190	210
DR cons	TCCTGGAGCAGAAGCGGGCCGCGGTGGACACCTACTGCAGACACAA	
DR 4 βIII	-----G-G-----A-----T-----	
	* * *	
DRw6b βI	-----G-G-----A-----T-----	
DRw6a βI	-----AG-CG-----C-----	
DR 5 βI	-----AG-C-G-----C-----	
DR 3 βI	-----G-CG-----A-----	
	* * *	
DRw6a βIII	-----G-CG-----AT-----	

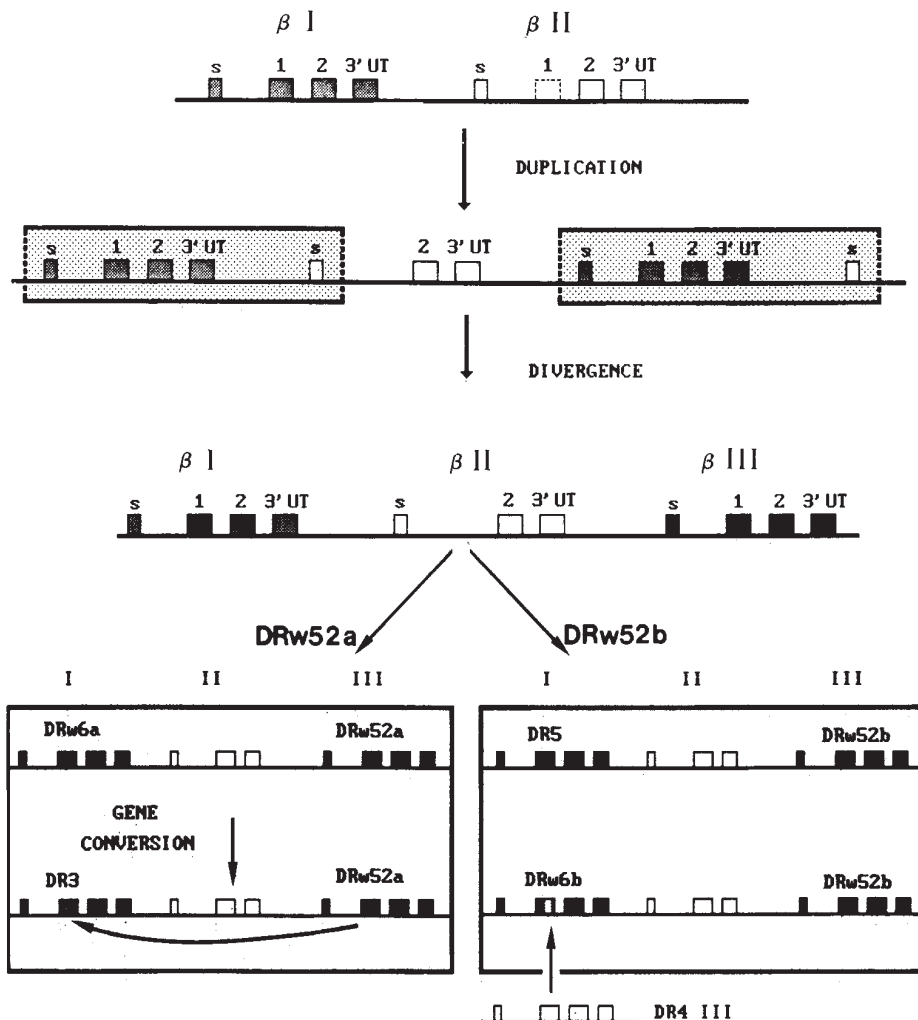
Fig. 3 An example of a possible interchromosomal gene conversion between distantly related haplotypes. The asterisks mark positions common to the *DR6b* βI - and *DR4* βIII -chains. The *DR3* conversion event is also marked. The *DR4* βIII sequence is derived from a Charon 30 phage library of the *DR4*, *DRw6* cell line mentioned previously.

gene conversion event was observed (Fig. 3). This region, which is similar in *DRw6a* βI and *DR5* βI , is very different in the βI locus of *DRw6b*. Surprisingly, it is identical to the same region in the βIII locus of *DR4*, a haplotype from a distantly related supertypic group, *DRw53*. It is thus likely that the *DRw6b* βI locus has resulted from a gene conversion between an ancestral gene similar to *DRw6b* βI as recipient and the βIII locus of the *DR4* haplotype which acted as donor.

This region of high allelic polymorphism, where we observe examples of gene conversion, is the same as that involved in the *I-E* to *I-A* conversion event giving rise to the *I-A^{bm12}* mutation in mice¹⁸⁻²⁰. Furthermore, a micropolymorphism reported in the *DR4* haplotype in this same DNA segment²¹, correlating with changes in alloreactive T-cell recognition, can be attributed to a gene conversion of the *DR4* βI locus, with the *DR4* βIII locus acting as donor. An interesting observation is that in all these cases the donor locus is 3' of the target locus with respect to the direction of transcription. Together, these results suggest a conversion hotspot responsible for this hyper-variable region. The presence of this hotspot implies either that conversion occurs here preferentially for physical reasons or that conversion is random but selection has favoured those events occurring in this region.

Concerning the less polymorphic *DRBIII* locus, nucleotide sequence comparison defines two groups of haplotypes, *DRw52a* and *DRw52b*, whose *DRBIII* loci differ by 11 bp. The first group, *DRw52a*, includes some *DR3* and some *DRw6* haplotypes (Fig. 1b). The second group, *DRw52b*, includes other *DR3* and *DRw6* haplotypes (Fig. 1b and unpublished results) as well as *DR5* (B. Schwartz, personal communication). Hybridization with appropriately chosen oligonucleotides has

Fig. 4 Evolution of the *DRw52* group of haplotypes. Restriction map homology around the βI and βIII loci indicates a recent duplication event (boxed and shaded). The duplicated loci then diverged, resulting in the present locus *I* and locus *III* (change in shading intensity). The *DRBII* locus lacks a first-domain encoding exon. The deletion of this exon may have followed or been part of the duplication event. Divergence at the βIII locus divides the *DRw52* family into two groups (*a* and *b*). Within group *a*, the *DR3* haplotype resulted from the gene conversion of the *DRw6a* βI locus (arrow). The *DRw6a* βI locus itself resembles the ancestral gene. The *DRBIII* gene in both haplotypes is identical. In group *b*, the *DRw6b* haplotype resulted, in part, from an interchromosomal gene conversion, with the *DR4* βIII locus acting as donor (arrow). The *DR5* βI locus resembles the ancestor. The βIII loci are almost identical. The loci are numbered in the direction of transcription.



allowed us to type unknown individuals as *DRw52a* or *DRw52b* (ref. 22 and unpublished) in a sizeable number of normal and diseased individuals. This form of analysis can be of considerable use in phylogenetic studies of human populations.

Taken together, these results can account for the evolution of the *DR* genes in the *DRw52* supertypic group (Fig. 4). The ancestral features of this family are the relatively recent duplication of the βI locus and the silencing of the βII locus by deletion of the first domain encoding exon¹¹. The duplicated *DRBI* loci then diverged into βI and βIII . Further divergence resulted in a branching into two lineages (*DRw52a* and *DRw52b*) based on common alleles at the less polymorphic locus, *DRBIII*. In the *DRw52a* group, the *DRw6a* haplotype gave rise to the *DR3* specificity by the gene conversion described here. The *DRw6b* haplotype was probably involved in an interchromosomal gene conversion with the *DR4 BI* locus acting as donor.

This analysis provides a framework for assigning serological specificities to the products of the different loci of the *DRw52* haplotypes. Allelic differences in the product of the *DRBIII* locus split *DRw52* into *a* and *b*. It has already been shown that this locus encodes the *DRw52* specificity for the case of the *DRw6b* haplotype²⁵. We propose that the distinct epitopes *DRw52a* and *DRw52b* (Fig. 4) will correspond to serological and T-cell specificities. In addition to the product of locus *DRBIII*, each haplotype obviously also expresses the product of their βI locus, which determines the fine *DR* specificity.

The data described here represent an example of relatively rapid evolution of a multigene family in which the loci appear to diverge at different rates following a duplication event. The time of divergence may be estimated by analysing this group of haplotypes in other geographical (non-European) groups whose migratory patterns are known.

The divergence in this gene family is generated in part by gene conversion. Since this mechanism can involve the transfer of preselected epitopes, the resulting additional polymorphism is frequently maintained in the population. Therefore, even if gene conversion is a relatively rare event, it can play a major role in the generation of polymorphism by producing functionally effective variants. It is generally thought that this polymorphism confers a selective advantage to a population in terms of its ability to cope with various pathogens. A genetic system with multiple loci undergoing conversion events could regenerate polymorphism in populations which have undergone bottlenecks due to migration or adverse environmental factors.

We thank Chantal Mattmann and Madeleine Zufferey for technical assistance, Dr B. Schwartz for providing sequence data for *DR5* prior to publication and Dr J. J. van Rood for comments on the manuscript. This work was supported by a grant from the Swiss National Fund for Scientific Research.

Note added in proof: From a recent analysis of micropolymorphism of HLA-*DR4* (ref. 26), we propose that one *DR4BI* allele, *Dw10*, has arisen by a gene conversion, with *DRw6aBI* acting as donor.

Received 17 January; accepted 4 April 1986.

- Schwartz, R. A. *Rev. Immun.* **3**, 237-261 (1985).
- Nagy, Z. *et al. Immun. Rev.* **60**, 61-83 (1981).
- Kaufman, J. *et al. Cell* **36**, 1-13 (1984).
- Kaufman, J. F. & Strominger, J. L. *Nature* **297**, 694-697 (1982).
- Long, E. O. *et al. EMBO J.* **2**, 389-394 (1983).
- Gorski, J. *et al. in UCLA Symp.* **18**, 47-56 (1984).
- Speis, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 5165-5169 (1985).
- Lee, J., Trowsdale, J. & Bodmer, W. F. *Proc. natn. Acad. Sci. U.S.A.* **79**, 545-549 (1982).
- Korman, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 1844-1848 (1982).
- Wake, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **79**, 6979-6983 (1982).
- Rollini, P., Mach, B. & Gorski, J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7197-7201 (1985).
- Tanigaki, N. & Tosi, R. *Immun. Rev.* **66**, 5-37 (1982).
- Angelini, G. *et al. in Histocompatibility Testing* (eds Albert, E. D., Baur, M. P. & Mayr, W. R.) 582-584 (Springer, New York, 1984).
- Hui, K. M. *et al. in Histocompatibility Testing* (eds Albert, E. D., Baur, M. P. & Mayr, W. R.) 590-594 (Springer, New York, 1984).
- Tiebert, V. L. *et al. J. biol. Chem.* **261**, 2738-2742 (1986).
- Kourilsky, P. *Biochimie* **65**, 85-93 (1983).
- Grosse-Wilde, H., Doxiadis, I. & Brandt, H. *in Histocompatibility Testing* (eds Albert, E. D., Baur, M. P. & Mayr, W. R.) 249-264 (Springer, New York, 1984).
- Denaro, M. *et al. EMBO J.* **3**, 2029-2032 (1984).

- Mengle-Gaw, L. *et al. J. exp. Med.* **160**, 1184-1194 (1984).
- Widera, G. & Flavell, R. A. *EMBO J.* **3**, 1221-1225 (1984).
- Cairns, J. S. *et al. Nature* **317**, 166-168 (1985).
- Angelini, G. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Gorski, J. *et al. J. exp. Med.* **162**, 105-116 (1985).
- Gregersen, P. K. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 2642-2646 (1986).

Hepatitis B virus DNA integration in a sequence homologous to *v-erb-A* and steroid receptor genes in a hepatocellular carcinoma

Anne Dejean*, Lydie Bougueleret†, Karl-Heinz Grzeschik‡ & Pierre Tiollais*

* Unité de Recombinaison et Expression Génétique (INSERM U.163, CNRS UA.271) and † Unité d'Informatique Scientifique, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris Cedex 15, France

‡ Institut für Humangenetik der Westfälischen Wilhelms-Universität, 4400 Münster, FRG

Hepatitis B virus (HBV) is clearly involved in the aetiology of human hepatocellular carcinoma (HCC)¹ and the finding of HBV DNA integration into human liver DNA in almost all HCCs studied²⁻⁷ suggested that these integrated viral sequences may be involved in liver oncogenesis. Several HBV integrations in different HCCs^{8,9} and HCC-derived cell lines¹⁰⁻¹⁴ have been analysed after molecular cloning without revealing any obvious role for HBV. From a comparison of a HBV integration site present in a particular HCC⁸ with the corresponding unoccupied site in the non-tumorous tissue of the same liver, we now report that HBV integration places the viral sequence next to a liver cell sequence which bears a striking resemblance to both an oncogene (*v-erb-A*) and the supposed DNA-binding domain of the human glucocorticoid receptor and human oestrogen receptor genes. We suggest that this gene, usually silent or transcribed at a very low level in normal hepatocytes, becomes inappropriately expressed as a consequence of HBV integration, thus contributing to the cell transformation.

We have previously reported the molecular cloning of the single integrated viral sequence present in the liver tumorous nodule of patient D and we have determined the sequences of the cellular-viral junctions⁸. The viral insertion was a continuous subgenomic fragment 1.4 kilobases (kb) long (Fig. 1a) containing the cohesive-end region, gene C and the beginning of gene *pre-S1*. We therefore used the 1.1-kb and 5.8-kb *HindIII* cellular fragments and the 1.8-kb *EcoRI* host-viral fragment (respectively referred to as LT (left tumour), RT (right tumour) and MT (medium tumour); Fig. 1a) to isolate the unoccupied site from a λ phage library of DNA extracted from the non-tumorous liver of patient D. This part of the liver did not seem to contain any integrated HBV sequences. Seven overlapping clones, hybridizing to one and/or the other of the three probes, were isolated and represented 32 kb of cellular DNA at the unoccupied site (Fig. 1b). Southern blots of restriction digests of the seven clones using total human DNA as a probe showed that the host sequence at the viral insertion site corresponds mainly to unique sequence DNA (Fig. 1a, b, solid bars). Comparison between the restriction maps of the unoccupied site (Fig. 1b) and the integrated site (Fig. 1a) did not reveal any major genomic rearrangements in the cellular DNA. Integration took place within a small *EcoRI* fragment of 400 base pairs (bp) which we subcloned and refer to as MNT (medium non-tumour) (Fig. 1b).

To investigate whether HBV became integrated in the vicinity of a cellular gene in the human genome, we determined the nucleotide sequence¹⁵ of the normal allele. This sequence

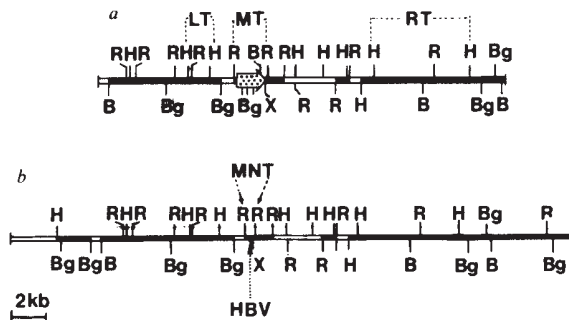


Fig. 1 *a*, Restriction map of the occupied site. This has previously been isolated from a library of cellular DNA extracted from the tumorous part of the liver of patient D⁸. The stippled region represents the 1.4-kb integrated HBV sequence. The arrowhead denotes the same orientation as the viral (+) strand³⁰. Solid bars denote unique sequences and open bars regions containing repetitive cellular sequences. Restriction sites are: R, *EcoRI*; B, *BamHI*; H, *HindIII*; X, *XhoI*. The cloned 1.1-kb *HindIII*, 1.8-kb *EcoRI* and 5.8-kb *HindIII* cellular fragments are referred to as LT, MT and RT. *b*, Restriction map of the unoccupied site. The cloning experiment in which partial *BclI* digests of cellular DNA were cloned into λ L47.1 was as described previously¹¹. A library, made with cellular DNA from the non-tumorous part of the liver of patient D, was screened with LT, MT and RT as probes. Seven positive overlapping clones were analysed, representing 32 kb of cellular DNA. The cloned 400-bp *EcoRI* cellular fragment is referred to as MNT. The site of HBV integration within MNT is indicated by an arrow.

revealed the presence of an open reading frame (ORF) of 519 nucleotides which was interrupted in the middle by the viral insertion (Fig. 2). Since methionine codons were present only at the 3' end of the ORF, this region can only correspond to a single exon from a split gene. Two acceptor-like (A-L1, A-L2) and two donor-like (D-L1, D-L2) consensus sequences of splice junctions¹⁶ could be localized, respectively, at the 5' end and the 3' end of the ORF (Fig. 2). A computer-assisted analysis¹⁷ showed that the region between A-L2 and D-L1 had a strong probability of being an exon. A search for related sequences in the NBRF protein data bank¹⁸ identified a remarkable homology (22 identities out of 49) between the translation product of the exon-like region of the ORF (amino-acid residues 69–117) and the amino-terminal region of the *v-erb-A* oncogene product (residues 8–58)¹⁹ (Fig. 3). Moreover, a significant homology has recently been reported between the *v-erb-A* protein and either the human glucocorticoid receptor (hGR)²⁰ or the human oestrogen receptor (ER)²¹. The alignment of the amino-acid sequence of hGR and ER with the exon-like protein product showed, in both cases, 19 identities out of 49 (Fig. 3). As already observed for *v-erb-A*, hGR²⁰ and ER²¹, the identity became greater beyond amino acid 96 of the ORF, which corresponds to a cysteine-rich region. In particular, the four cysteines conserved between *v-erb-A*, hGR and ER are present at the same position in the ORF. Although the homology continues for *v-erb-A*, hGR and ER—all derived from cDNA clones—no significant homology was found beyond residue 117 for the ORF. However, since the D-L1 sequence (next to residue 117) is strictly identical to the consensus sequence of a donor site, it could thus correspond to an exon/intron boundary.

The integration of HBV sequences interrupted the cellular open reading frame and generated a microdeletion of 7–12 bp (boxed in Fig. 4). This minor rearrangement provides evidence that the situation we are studying in patient D is probably very near the initial integration event. In addition to the microdeletion, the viral integration—interrupting the cellular ORF—generated a new viral–host hybrid sequence such that the first 29 codons of the viral *pre-S1* gene became fused and in phase with the last 28 codons of the cellular exon (Fig. 4). Remarkably,

```

CTTAGTTAGACAAGTCACTTCTCTCTTTGAGTTTCTGTTCTGTTGGCTAAATTTG
1
TGATAAAACCATGGCCCTGAGAGTTCTGTGTGAGCTGTGGACTGCAAGTGAGAGATGA
50
*** Asp Cys Phe Val Asn Arg Val Pro His Gln Cys Gln Cys His Cys Trp Ala
TGACACTGCTTTGTAAACCGTAGAGTACCACATCAATGTCAGTGTCAACACTGCTGGGCC
120
20
Tyr Asn Asp Ser Cys Phe Tyr Glu Pro Ile Leu Ala Ser Val Ile Ala Glu Phe Gly Ala
TCAATGACAGCTGTTTTATGACCCATTCTTCTGTAGTGTATTGCTGAATTTGGGCCA
180
--- A-L1 ---
40
Glu Ala Tyr Glu Phe Thr Val Glu Lys Ser Asn Glu Leu Lys Tyr Ile Phe Ser Phe Phe
GAACATATGAATTCAGTGTGAAAAAGTAATGAACATAAATACATTTCTCTTTTTC
240
60
Ser Leu Tyr Pro Ile Ser Ile Ile Ala Ile Glu Thr Gln Ser Thr Ser Ser Glu Glu Leu
TCCCTGACCCCATCTCTATTATGCAATTGAAACACAGAGCACCAGCTGTGAGAACTC
300
--- A-L2 ---
80
Val Pro Ser Pro Pro Ser Pro Leu Pro Pro Pro Arg Val Tyr Lys Pro Cys Phe Val Cys
GTCCCAAGCCGCCCATCTCCACTTCTGCTCCCTCGAGTGTACAAACCTGCTTCTGCTGC
360
HBV integration
100
Gln Asp Lys Ser Ser Gly Tyr His Tyr Gly Val Ser Ala Cys Glu Gly Cys Lys Val Ser
CAGGACAAATCATCAGGATCACCATATGGGTCAGCGCTGTGAGGATCTAAGGTGACT
420
--- D-L1 ---
120
Ile His Thr Ser Val Pro Asp Glu Leu Ser Phe Ser Met Tyr Phe Met Glu Val Thr Tyr
ATTACACTCTCTGCTGATGAACTCTCATCTTCATCTGACTTTATGGAGGTGACCTAC
480
140
Pro Val Pro Lys Met Gly Trp Met Cys Asn Ile Thr Ser His Lys Cys Gly Glu Phe Ser
CCAGTTCAAAAATGGGCTGGATGTGTATACATCACCCTCCCAAGTGTGTGCAATCAGT
540
--- D-L2 ---
160
Tyr Ile Gln Trp Phe Leu Leu Leu Pro Tyr Ile Ser Leu Val ***
TATATTCACTGGTTTTATTACTTCTTATATATCTTGGTTTAAATTCAGAACTTCTG
600
660
CAATGACTGATGAGGCTCAAGGTTCTAAGGCCCTTAAATACCCCAACATAAATTC
720
CTTTAGATAATAAGAACTGCTAGGCCACTGTGACTGC
720

```

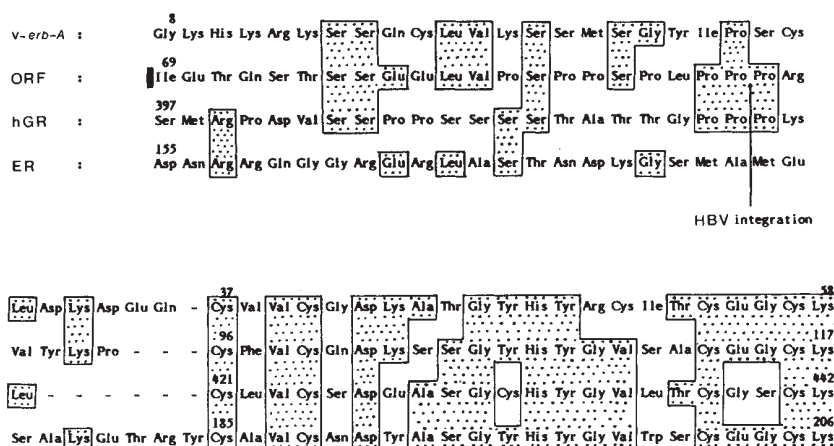
Fig. 2 Nucleotide sequence of the unoccupied site. Nucleotides are numbered at the left side. The deduced amino-acid sequence of the 519-bp open reading frame is shown above the nucleotide sequence. The amino-acid sequence is numbered from the first codon of the ORF. A large number of splice junction sequences have been reported¹⁶. The compilation of the data supports the consensus (C)₁₁N^CAG/G for acceptors and the consensus CAG/GT^AAGT for donors. The two acceptor-like (A-L1 and A-L2) and donor-like (D-L1 and D-L2) sequences are underlined. The site of HBV integration in the middle of ORF is indicated by an arrow. The cloned 3.4-kb *HindIII* fragment, encompassing the unintegrated site in the normal allele, was sonicated, treated with the Klenow fragment of DNA polymerase plus deoxyribonucleotides (2 h, 15 °C) and fractionated by agarose gel electrophoresis. Fragments of 400–700 bp were excised and electroeluted. DNA was ethanol-precipitated, ligated to dephosphorylated *SmaI*-cleaved M13 mp8 replication form DNA and transfected into *Escherichia coli* strain TG-1 by the high-efficiency technique of Hanahan³¹. Recombinant clones were detected by plaque hybridization using the MNT (Fig. 1b) subclone DNA as a probe. Single-stranded templates were prepared from plaques exhibiting positive hybridization signals and were sequenced by the dideoxy chain termination procedure¹⁵ using buffer gradient gels³².

the viral genome became integrated a few nucleotides upstream from the most conserved Cys-rich portion of the ORF (Fig. 3), maintaining the integrity of this region.

Using a panel of 17 mouse–human and Chinese hamster–human somatic cell hybrid DNAs^{22,23}, we localized the ORF to chromosome 3 (data not shown), while the *c-erb-A* oncogene²⁴, hGR²⁵ and ER²⁶ have been mapped, respectively, to human chromosomes 17, 5 and 6. In preliminary experiments, we hybridized the MNT probe, encompassing the exon-like region of ORF, to a Northern blot of polyadenylated RNAs extracted from five human livers, but found no detectable transcripts. A large number of human fetal and adult tissues will have to be tested similarly to reveal any active transcription of this region.

The conserved Cys-rich region which extends over 60 amino-acid residues in *v-erb-A* protein, hGR and ER is thought to include the DNA-binding domain of the molecule^{20,21}. We can thus speculate that the corresponding homologous region of ORF, truncated by the exon–intron boundary, is part of a cellular gene that shares a common functional domain with hGR, ER

Fig. 3 Amino-acid sequence alignment of the *v-erb-A* oncogene protein¹⁹, the translation product of the exon-like region of ORF, the human glucocorticoid receptor (hGR)²⁰ and the oestrogen receptor (ER)²¹. The limits of the exon-like region of ORF, defined by A-L2 and D-L1 boundaries, are indicated by rectangles. To predict the location of exon-like regions, we used the discriminating program PREDICTOR¹⁷. Two subsets of the GenBank data library, containing either only exon or only intron sequences, were taken as reference pool. The program PROBE3-EXPLOR3 (ref. 18), allowing the search for ambiguous nucleic or peptidic patterns, was used to screen both the NBRF (proteins) and the GenBank (DNA) data banks. These programs were run on a MV8000 32-bit minicomputer. Amino-acid residues 69–117 from the ORF were aligned with amino-acid residues 8–58 from p75^{erb-A}, residues 397–442 from hGR and residues 115–206 from ER. Identical residues are boxed and gaps are indicated by dashes. The HBV integration site, upstream from the cysteine-rich region, is indicated.



and *v-erb-A* gene products and which could exert a transcriptional regulatory function on specific genes.

Although the way in which HBV participates in the formation of a liver cancer is unknown, the experiments reported here could promote our understanding of one possible mechanism of HBV carcinogenesis. In patient D the viral integration, interrupting the exon-like region (Fig. 2), created a chimaeric viral-host open reading frame (Fig. 4). The HBV insertion took place a few nucleotides upstream from the beginning of the putative DNA-binding domain. Since a viral promoter has been defined by *in vitro* transcription approximately 30 nucleotides upstream from the initiator codon of the *pre-S1* gene²⁷, we suggest that, in the tumorous part of the liver, a readthrough transcription occurred from the viral promoter. Although protein or RNA from the tumour is no longer available to test this hypothesis, it is most probable that inappropriate activation of the putative gene as a consequence of HBV integration resulted in expression of a truncated protein at greater levels than that of the native protein. This protein could participate directly in the subsequent cell transformation.

Several arguments suggest that hormonal factors are involved in human hepatocarcinogenesis. The incidence of HCC is three- to sixfold greater in males¹, and the use of oral contraceptives in females is associated with the development of hepatic adenoma²⁸. Moreover, the ability of oestrogenic hormones to function as promoters of neoplastic development in rat liver has been clearly demonstrated²⁹. The finding that HBV sequences have become integrated into a putative cellular gene sharing homology with the steroid receptor genes is therefore intriguing and suggests that, in some cases, hormonal and HBV carcinogenesis may be directly related.

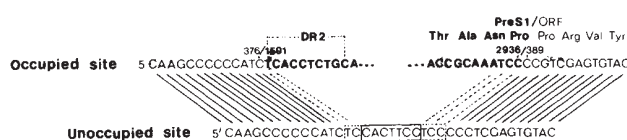


Fig. 4 DNA sequences at the HBV integration site. The sequences of the left and right host-viral DNA junctions at the occupied site (upper sequence) are compared with the human DNA sequence at the unoccupied site (lower sequence). The bold-face letters indicate the viral sequence. Nucleotides of the ORF and the HBV genome³³ are numbered. Homologous nucleotides between the two sequences are indicated by sloping lines. The 7-bp CACTTCC present in the normal allele and deleted after the viral integration in the occupied site is boxed. Because HBV DNA and cellular DNA shared a 2-bp and 3-bp sequence homology at a point coincident, respectively, with the left and right host-viral junctions (dashed lines), the deleted fragment could be up to 12 bp long. The DR2 copy of the 11-bp viral direct repeat specifically involved in HBV integration⁸ is indicated. The putative chimaeric protein, generated by the HBV inversion, between the first 29 amino acids of the viral *pre-S1* gene product and the last 28 amino acids of the cellular exon protein product is partially represented at the fusion site.

We thank Drs Christian Bréchet, Pierre Sonigo, Simon Wain-Hobson and Jean Weissenbach for advice and discussion and Louise-Marie Da for typing the manuscript. This work was supported by NIH grant CA97300-02, the Economic European Community and the Faculté de Médecine Lariboisière Saint-Louis (University of Paris VII).

Received 5 February; accepted 16 April 1986.

- Beasley, R. P. & Hwang, L. Y. in *Viral Hepatitis and Liver Disease* (eds Vyas, G. N., Dienstag, J. L. & Hoofnagle, J. H.) 209–214 (Grune & Stratton, New York, 1984).
- Chakraborty, P. R., Ruiz-Opazo, N., Shouval, D. & Shafritz, D. A. *Nature* **286**, 531–533 (1980).
- Bréchet, C., Pourcel, C., Louise, A., Rain, B. & Tiollais, P. *Nature* **286**, 533–535 (1980).
- Edman, J. C., Gray, P., Valenzuela, P., Rall, L. B. & Rutter, W. J. *Nature* **286**, 535–538 (1980).
- Bréchet, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 3906–3910 (1981).
- Shafritz, D. A., Shouval, D., Sherman, H. I., Hadziyannis, S. J. & Kew, M. C. *New Engl. J. Med.* **305**, 1067–1073 (1981).
- Chen, D. S., Hoyer, B. H., Nelson, J., Purcell, R. H. & Gerin, J. L. *Hepatology* **2**, 42S–45S (1982).
- Dejean, A., Sonigo, P., Wain-Hobson, S. & Tiollais, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 5350–5354 (1984).
- Rogier, C. E. *et al. Science* **230**, 319–322 (1985).
- Koshy, R. *et al. Cell* **24**, 215–223 (1983).
- Dejean, A., Bréchet, C., Tiollais, P. & Wain-Hobson, S. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2505–2509 (1983).
- Shaul, Y. *et al. J. Virol.* **51**, 776–787 (1984).
- Mizusawa, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 208–212 (1985).

- Yaginuma, K., Kobayashi, M., Yoshida, E. & Koike, K. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4458–4462 (1985).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Mount, S. M. *Nucleic Acids Res.* **10**, 459–472 (1982).
- Claverie, J. M. & Bougueleret, L. *Nucleic Acids Res.* **14** (in the press).
- Claverie, J. M. & Sauvaget, I. *Cabios* **1**, 95–104 (1985).
- Debuire, B. *et al. Science* **224**, 1456–1459 (1984).
- Weinberger, C., Hollenberg, S. M., Rosenfeld, M. G. & Evans, R. M. *Nature* **318**, 670–672 (1985).
- Green, S. *et al. Nature* **320**, 134–139 (1986).
- Johannsmann, R., Hellkuhl, B. & Grzeschik, K. H. *Hum. Genet.* **56**, 361–363 (1981).
- Huerre, C. *et al. Ann. Genet.* **27**, 5–10 (1984).
- Spurr, N. K. *et al. EMBO J.* **3**, 159–163 (1984).
- Hollenberg, S. M. *et al. Nature* **318**, 635–641 (1985).
- Walter, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 7889–7893 (1985).
- Rall, L. B., Standing, D. N., Laub, O. & Rutter, W. J. *Molec. cell. Biol.* **3**, 1766–1773 (1983).
- Edmonson, M. A., Henderson, B. & Benton, B. *New Engl. J. Med.* **294**, 470 (1976).
- Wanless, I. R. & Medline, A. *Lab. Invest.* **46**, 313–320 (1982).
- Tiollais, P., Pourcel, C. & Dejean, A. *Nature* **317**, 489–495 (1985).
- Hanahan, D. *J. molec. Biol.* **166**, 557–580 (1983).
- Biggin, M. D., Gibson, T. J. & Hong, G. F. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3963–3965 (1983).
- Galibert, F., Mandart, E., Fitoussi, F., Tiollais, P. & Charnay, P. *Nature* **281**, 646–650 (1979).

Analysis of deletions in DNA from patients with Becker and Duchenne muscular dystrophy

Louis M. Kunkel and co-authors*

Division of Genetics, The Children's Hospital, Boston, Massachusetts 02115, USA

Duchenne muscular dystrophy (DMD) is an X-linked recessive genetic disorder for which the biochemical defect is as yet unknown. Recently, two cloned segments of human X-chromosome DNA have been described which detect structural alterations within or near the genetic locus responsible for the disorder^{1,2}. Both of these cloned segments were described as tightly linked to the locus and were capable of detecting deletions in the DNA of boys affected with DMD. In an attempt to determine more precisely the occurrence of these deletions within a large population of DMD patients and the accuracy of one of the segments, DXS164 (pERT87), in determining the inheritance of the DMD X chromosome, the subclones 1, 8 and 15 were made available to many investigators throughout the world. Here we describe the combined results of more than 20 research laboratories with respect to the occurrence of deletions at the DXS164 locus in DNA samples isolated from patients with DMD and Becker muscular dystrophy (BMD). The results indicate that the DXS164 locus apparently recombines with DMD 5% of the time, but is probably located between independent sites of mutation which yield DMD. The breakpoints of some deletions are delineated within the DXS164 locus, and it is evident that the deletions at the DMD locus are frequent and extremely large.

The previously described deletions of the DXS164 region and the XJ-1.1 junction clone^{1,2} were assumed to be large, for neither set of cloned segments exhibits any common sequence overlap and yet both cloned segments are absent from most of the same deletion DNA samples². To increase the possibility that a particular deletion would exhibit a break within cloned DNA, the DXS164 region was expanded further by chromosome walking³ in human genomic libraries constructed in the phage vector EMBL-3 (ref. 4). Following five bidirectional walks, a 137-kilobase (kb) contiguous stretch of DNA was obtained from the DXS164 region. All DNA segments obtained from the 137 kb of DNA were subcloned in plasmid vectors and segments of unique sequence were identified for the entire length. The two previously described subclones (pERT87-1 and -8; ref. 1) and a new unique-sequence segment (pERT87-15) which also detects restriction fragment length polymorphisms (RFLPs; ref. 5) were sent to other investigators interested in DMD.

The three DXS164 subclones, which are spaced over a 50-kb

section, were tested for deletion by hybridization against the DNA isolated from males exhibiting the DMD and BMD phenotype. The results of this analysis for 1,346 males are presented in Table 1. As the results were obtained in many different laboratories, they are presented separately for each laboratory. Of all DMD and BMD males tested, 6.5% show deletions at the DXS164 locus, a slightly lower percentage than that found previously¹. A difference was observed in the incidence of deletions between DMD males with a clear family history of the disease (8.3%) and those with no family history (5.8%). Such a difference might be a reflection of the incidence of deletion mutation in maternal and paternal meiosis. Any true sporadic case of DMD where the mother is not a carrier of the disease represents a mutation which must occur only in female meiosis with no paternal contribution. Familial cases would be assumed to have a contribution of both male and female meiotic deletion events. Thus, if the frequency of deletion mutations at the DMD locus were nearly equal in males and females, one might assume that a higher incidence of deletions, as presented in Table 1, would be observed in familial cases. Prediction of affected individuals in families segregating a deletion which is assumed to be the primary genetic cause of the disease (of more than 150 normal boys, no deletions were observed) should be highly accurate.

As indicated in Table 1, two boys with BMD were found to have a deletion of the DXS164 region. Linkage analysis has shown that the locus for BMD is localized near, or is a potential allele of, DMD⁶⁻⁹. The finding of deletions in the DNAs of BMD males which overlap with deletions found in DMD males suggests that if they are two separate loci they are indeed close to one another. Alternatively, if BMD and DMD are caused by different alleles at the same locus, then the milder BMD phenotype might be expected to be the result of a low-level expression of the DMD/BMD gene product. If this were the case, the BMD deletions might not involve DNA sequences that are absolutely necessary for the expression of the DMD/BMD gene product.

The three DNA probes from the DXS164 region that were distributed to other investigators each recognize the informative RFLP alleles given in Table 2. The separate RFLP alleles of pERT87-1 and -8 exhibit a degree of disequilibrium and are only informative in ~60% of women tested, despite there being many different enzyme-defined loci. The three separate loci of pERT87-15 defined by different enzymes, although in some degree of disequilibrium with each other, are closer to equilibrium with the pERT87-1 and -8 loci. When the three enzyme-defined loci of the pERT87-15 probe were used in combination with the *Bst*XI locus of pERT87-8 and the *Xmn*I locus of pERT87-1, 25 of 28 (89%) unrelated women at risk for DMD or BMD were observed to be informative for linkage between one or more of the RFLP-detecting loci and the disease locus.

* J. F. Hejtmancik & C. Th. Caskey, Institute of Molecular Genetics, Baylor College of Medicine, Houston, Texas 77030, USA; A. Speer, Central Institute of Molecular Biology, Academy of Science DDR, 1115 Berlin, DDR; A. P. Monaco, W. Middlesworth, C. A. Colletti, C. Bertelson & U. Müller (Division of Genetics), M. Bresnan (Department of Neurology), F. Shapiro (Department of Orthopedics), U. Tantravahi, J. Speer & S. A. Latt (Division of Genetics, DMD Diagnostic Laboratory), The Children's Hospital, Boston, Massachusetts 02115, USA; R. Bartlett, M. A. Pericak-Vance & A. D. Roses, Division of Neurology, Duke University Medical Center, Durham, North Carolina 27710, USA; M. W. Thompson, P. N. Ray & R. G. Worton, Department of Genetics, Hospital for Sick Children, Toronto, Ontario, Canada M5G 1X8; K. H. Fischbeck, Neurology Department, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; P. Gallano, M. Coulon, C. Duros, J. Boue & C. Junien, INSERM 473, Chateau de Longchamp, 75016 Paris, France; J. Chelly, G. Hamard, M. Jeanpierre, M. Lambert & J.-C. Kaplan, Institut de Pathologie Moléculaire, INSERM 129, Chu Cochin, 75014 Paris, France; A. Emery, Medical School, Teviot Place, Edinburgh EH18 9GA, UK; H. Dorkins, S. McGlade & K. E. Davies, Nuffield Department of Clinical Medicine, John Radcliffe Hospital, Oxford OX3 9DU, UK; C. Boehm, Department of Pediatrics, Johns Hopkins University, Baltimore, Maryland 21205, USA; B. Arveiler & C. Lemaire, Laboratoire de Génétique Moléculaire des Eucaryotes, CNRS and INSERM U184, Faculté de Médecine, Strasbourg, France; G. J. Morgan & M. J. Denton, Department of Pathology, Prince of Wales Hospital, Sydney, New South Wales 2031, Australia; J. Amos, Eunice Kennedy Shriver Center, Waltham, Massachusetts 02254, USA; M. Bobrow, F. Benham, E. Boswinkel, C. Cole, V. Dubowitz, K. Hart, S. Hodgson, L. Johnson & A. Walker, Paediatric Research Unit, United Medical and Dental Schools of Guy's and St Thomas's Hospitals/Hammersmith Hospital, London SE1 9RT, UK; L. Roncuzzi, A. Ferlini, C. Nobile & G. Romeo, Laboratory of Molecular Genetics, Istituto G. Gaslini, 16148 Genova Quarto, and Laboratory of Genetics, Clinica Neurologica, Università di Bologna, Italy; D. E. Wilcox, N. A. Affara & M. A. Ferguson-Smith, Duncan Guthrie Institute of Medical Genetics, University of Glasgow, Yorkhill, Glasgow G3 8SJ, UK; M. Lindlof, H. Kaariainen & A. de la Chapelle, Department of Medical Genetics, University of Helsinki, 00290 Helsinki, Finland; V. Ionescu, Ch. Seaby & R. Ionescu, Department of Pediatrics, University of Iowa, Iowa City, Iowa 52242, USA; E. Bakker, G.-J.B. van Ommen & P. L. Pearson, Department of Human Genetics, Sylvius Laboratories, University of Leiden, The Netherlands; C. R. Greenberg, J. L. Hamerton & K. Wrogemann, Departments of Human Genetics and Biochemistry, University of Manitoba, Winnipeg, Manitoba, Canada R3E 0W3; R. A. Doherty, R. Polakowska, C. Hyser & S. Quirk, Division of Genetics and Dysmorphology, University of Rochester Medical Center, Rochester, New York 14642, USA; N. Thomas, Department of Medicine, University of Wales College of Medicine, Heath Park, Cardiff CF4 4XN, UK; J. F. Harper, B. T. Darras & U. Francke, Department of Genetics, Yale University School of Medicine, New Haven, Connecticut 06510, USA.

Table 1 Structural alterations detected at the DXS164 locus in DMD males

Contributing laboratory	Clinical description			Deletions		
	DMD-F	DMD-S	BMD	DMD-F	DMD-S	BMD
R. Bartlett, Durham*	11	9	2	3	1	0
M. Bobrow, London	75	50	24	7	2	1
T. Caskey, Houston	22	18	0	5	3	0
K. Davies, Oxford†	42	68	0	0	2	0
A. de la Chapelle, Helsinki	10	20	2	0	1	1
M. Denton, Sydney*	12	15	0	6	0	0
R. Doherty, Rochester	32	61	8	4	4	0
A. Emery, Edinburgh†	0	30	0	0	0	0
M. Ferguson-Smith, Glasgow	37	44	17	3	3	0
K. Fischbeck, Philadelphia*	25	27	4	1	1	0
U. Francke, New Haven	11	1	1	1	0	0
C. Greenberg, Manitoba	3	0	1	1	0	0
P. Harper, Cardiff	50	43	35	4	1	0
V. Ionasescu, Iowa City	16	21	5	2	1	0
C. Junien, Paris	51	2	0	1	1	0
C. Boehm, Baltimore	12	5	0	1	0	0
J. C. Kaplan, Paris	52	29	7	3	2	0
L. Kunkel, Boston*	14	29	0	1	3	0
J.-L. Mandel, Strasbourg	12	13	0	0	0	0
P. Pearson, Leiden	43	24	8	5	3	0
A. Read, Manchester†	22	32	0	2	2	0
G. Romeo, Genova	7	6	10	1	1	0
A. Speer, E. Berlin†	51	0	18	1	0	0
U. Tantravahi, Boston	8	1	0	2	0	0
R. Worton, Toronto*	32	3	3	0	1	0
Total	650	551	145	54	32	2
Percentage				8.3%	5.8%	1.4%
Overall total			1,346		88 (6.5%)	

The DNAs isolated from 1,346 DMD and BMD males from separate kindreds were cleaved with various restriction enzymes. Diagnosis of DMD and/or BMD was established by characteristic history, physical examination, increased CPK levels and, in many cases, a positive muscle biopsy¹⁹⁻²¹. Following separation of digested DNA by electrophoresis, the DNA samples were hybridized²² with one or more of the DXS164 subclones pERT87-1, pERT87-8 and pERT87-15. Not all samples were hybridized with all the subclones, but most were hybridized with pERT87-8. The investigator's laboratory in which the DNA isolations and, in most cases, the hybridizations were accomplished are indicated. The patients were divided into three categories: DMD-F, DMD males from a family in which there was more than one affected individual or, in a few cases, where the mother had an elevated CPK, but no clear family history of DMD; DMD-S, DMD males where there was no family history of the disease and the mother had a normal CPK value; BMD, males who exhibited the less severe X-linked myopathy. The number of boys studied within each of these three categories is shown for each laboratory. Deletions were defined as complete absence of hybridization for a DXS164 subclone and the number of individuals with absent fragments is given in each category. Those DMD males who exhibited deletions were representative of the clinical spectrum for DMD and included males who were mentally retarded as well as those not mentally retarded. Many DMD boys who were diagnosed as mentally retarded did not exhibit a deletion of DXS164 subclones. Those BMD males who exhibited deletion of DXS164 subclones were both on the more severe side of the BMD clinical spectrum, but both were wheelchair-bound later than a typical DMD patient: one was from a familial case of BMD and the other was a sporadic occurrence of BMD. Both are functioning well in their mid-twenties and one male is still able to drive a car.

* Most of these DNA samples were obtained from the individual listed and were sent to L.M.K. in Boston, where they were tested for deletion.

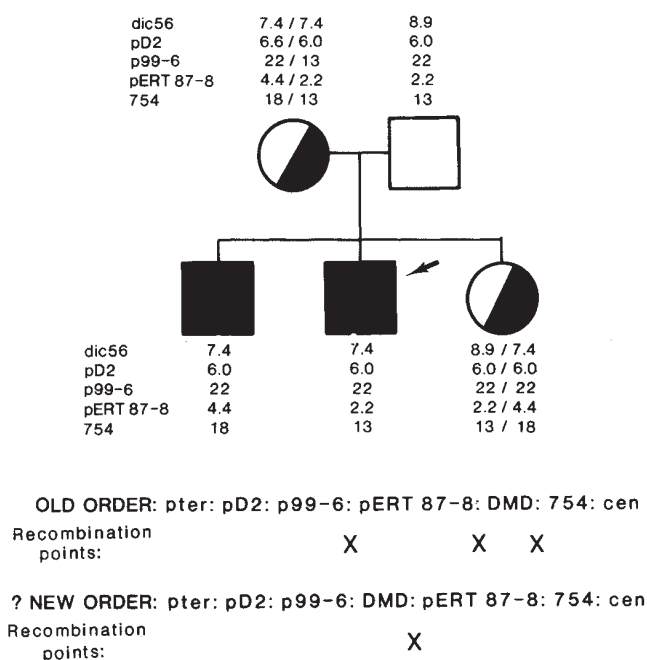
† Most of these DNA samples were sent to K.E.D. at Oxford, where they were tested for deletion. K.E.D. received other DNA samples from Dr Schwartz (Denmark) and Drs Ch. Coutelle, Spiegler, Herrmann and Szibor (FRG).

Segregation analysis of these RFLP-detecting loci in multiple families at risk for BMD and DMD has shown that recombinants are possible between DXS164 and the disease locus; one such recombinant is shown in Fig. 1. The family consists of two affected sons and a daughter whose level of creatine phosphokinase (CPK) and muscular hypertrophy indicate a carrier status. The boy indicated by an arrow in the figure has clearly received from his mother a 2.2-kb hybridizing *Bst*XI fragment for pERT87-8, whereas his affected brother and carrier sister have each received a 4.4-kb fragment. Thus, the X chromosome of the boy carrying the 2.2-kb fragment must be the result of a recombination event between DXS164 locus and DMD.

To determine more accurately the meiotic exchange points in this family and to minimize the possibility of sample error, the same family was also tested for segregation with other Xp cloned loci. The *Pst*I RFLP detected by the cloned probe 754 (DXS84; ref. 10) also exhibited recombination in the same individual, whereas the other loci, DXS41 (99-6) and DXS43 (D2)^{11,12}, appear to be non-recombinant. This is an unexpected result based on previous notions of gene order for these loci on the

X chromosome¹⁰⁻¹⁵. Physical mapping of these marker loci using panels of hybrid cells^{16,17} containing translocation chromosomes from two independent X; autosome translocations found in females with DMD localize DXS164 (ref. 14), DXS41 and DXS43 (refs 11, 12) on the pter side and DXS84 (refs 13, 17) on the centromere side of the translocation breakpoints. If the translocation exchange points in these two patients are in the DMD gene^{17,18}, the order of these loci would be as presented in Fig. 1 ('old order'). The recombination data of Fig. 1 are inconsistent with this order in this particular family. There must be a recombination event between DMD and DXS84(754), one between DMD and DXS164(pERT87-8), and another between DXS164(pERT87-8) and DXS41(99-6) for this latter marker to appear non-recombinant. The presumed order would require three meiotic exchanges within a region of previously closely linked markers^{9,15}. An alternative, and perhaps more likely explanation is that the DMD mutational site in this family maps between DXS41 and DXS164 and a single recombination (indicated at the bottom of Fig. 1) between DXS164 and the DMD mutation site would explain the results.

Fig. 1 Segregation of five Xp cloned loci in a DMD family. ○, Females; □, males; ■, DMD-affected individual; ◐, DMD carrier. DNA was isolated¹¹ from five members of a DMD kindred. Diagnosis of DMD was established as described in Table 1 legend, and the mother was judged to be a carrier of the disease as she exhibited an increased concentration of CPK^{19,20} and had given birth to two affected individuals. The mother, though, had no knowledge of any other case of DMD in her family. The daughter was judged to be a carrier as she had extremely elevated CPK (relative to other female carriers) and exhibited hypertrophy of the leg muscles, although no weakness. The DNA samples were cleaved with the various restriction enzymes appropriate for each of the cloned DNA probes. Gel electrophoresis, transfer and hybridization were as described previously¹¹. The different loci were tested by hybridization with the following cloned DNA fragments (given from the most distal): DXS143, dic56; DXS43, pD2; DXS41, p99-6; DXS164, pERT87-8; and DXS84, 754 (the most proximal). The various cloned segments had been physically mapped previously¹⁰⁻¹⁴. The numbers above and below each symbol indicate the size of hybridizing restriction fragments (alleles). Segregation of the locus DXS143 (ref. 23) is included to show that the samples from the non-recombinant boy and the father were not inverted. Two possible orders for the loci are indicated at the bottom of the figure: the first places the DMD locus at the point where the translocation breaks occur in two females with DMD; the second order (new order) represents an interpretation of the results obtained in the family study.



In the cases where DXS164 has been observed to recombine with DMD (eight known instances) and where DXS84 was also informative (five instances), both markers were recombinant. These results are consistent with the fact that the mutations giving rise to DMD in some families lie distal to DXS84, two translocation breaks in Xp21 (refs 16, 17) and the DXS164 locus. Thus, DXS164 may lie between these two translocation breaks and additional sites of mutation, all of which yield the phenotype of DMD and/or BMD. Alternative explanations for the linkage results could include a different X-linked locus with a phenotype very similar to DMD or new mutations which would appear as recombinants. Potentially consistent with the latter possibility is the observation that only one of the eight DXS164 recombinants observed occurs in the third generation or thereafter in a multi-generation DMD family.

For whatever reason, the exclusive use of DXS164 cloned probes as linked markers would lead to some mistakes when predicting the inheritance of the DMD X chromosome. This result has major implications for the prediction of the DMD genotype. Although the exact error rate is difficult to calculate (because only recombinants for other closely linked markers were tested for the segregation of DXS164 clones), best estimates can be made from the observation that approximately one-half of DXS84 recombinants (estimated at 10% recombination; ref. 9) are also DXS164 recombinants. This leads to the assumption that DXS164 will be in error in predicting the DMD X chromosome 5% of the time. It is strongly advised that the DXS164 locus be used in combination with other Xp21 flanking markers to construct an Xp21 haplotype. Given that the DXS164 locus seems to be localized between mutations found in different families which yield the DMD phenotype, DXS164 should not be considered as a flanking marker to DMD with other Xp21 loci. In a diagnostic situation, it would be unknown in which direction the particular mutation actually occurred relative to DXS164. Combined multi-locus information should improve the accuracy of genotype prediction in DMD families, assuming, of course, that the observed prediction errors for the DXS164 locus are caused by recombinants and not by new mutations or a different, as yet uncharacterized, locus.

The deletions observed in the DNAs of boys with DMD (and also BMD) provide the means to start to define important

segments of Xp21 which might be involved in the normal functioning of the DMD locus. Therefore, 53 of the 88 DNA samples which constituted deletions for the three DXS164 subclones distributed were tested for the absence or presence of additional subclones from the entire 137 kb of DNA from the DXS164 locus. A representative hybridization analysis is shown in Fig. 2 for five DXS164 subclones hybridized to nine different deletion DNA samples. The subclones are located within the DXS164 locus on either side of pERT87-1 and -8, and the five subclones combined can detect size differences between *Pst*I-hybridizing fragments over 45 kb of the DXS164 locus. Among the nine deletion DNA samples, five showed no DXS164-hybridizing *Pst*I fragment. The remaining four DNA samples exhibited some fragments but not others; these samples thus have breakpoints within the DXS164 locus, and one (Fig. 2, lane 3) of the breakpoints is preliminarily determined as an altered *Pst*I hybridizing fragment. By using additional subclones from the DXS164 locus, each breakpoint can be mapped more precisely to a specific subregion. Once localized to a specific region, the appropriate restriction enzyme can be chosen to best demonstrate the actual breakpoint (within 100–1,000 base pairs (bp)).

Table 2 RFLP-detecting subclones of DXS164

Subclone	Insert size (kb)	Enzyme	Allele size (kb)		Allele frequency	
			p	q	p	q
pERT87-1	1.3	<i>Bst</i> NI	3.1	2.5/0.6	0.63	0.37
		<i>Xmn</i> I	8.7	7.5	0.66	0.34
pERT87-8	1.3	<i>Bst</i> XI	4.4	2.2	0.6	0.4
		<i>Taq</i> I	2.7/1.1	3.8	0.71	0.29
pERT87-15	1.5	<i>Bam</i> HI	7.1/2.3	9.4	0.62	0.38
		<i>Taq</i> I	3.1	3.3	0.67	0.33
		<i>Xmn</i> I	1.6/1.2	2.8	0.68	0.32

The three DXS164 subclones are listed together with the sizes of their cloned human DNA inserts. The enzyme site variation detected by each probe is represented by the sizes of the hybridizing fragments and the respective frequencies of the common allele (p) and the rare allele (q). The allele frequencies were calculated in each case from the results obtained with over 75 X chromosomes.

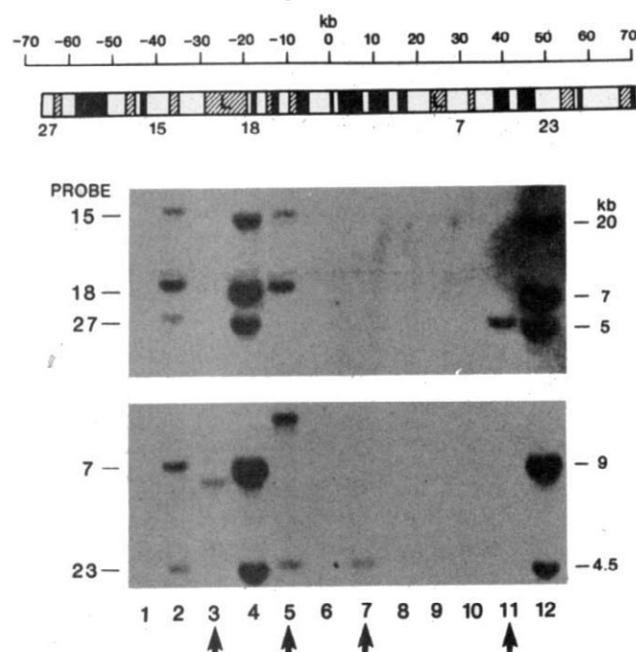


Fig. 2 Demonstration of four deletion breakpoints within the DXS164 locus. The DXS164 locus was expanded by chromosome 'walking'³ as described previously¹. A schematic drawing of 137 kb of contiguous human DNA is presented at the top. Solid bars represent regions of repeated DNA sequences; open boxes represent regions of unique sequence. Cross-hatched boxes represent moderately repeated sequences, and two copies of LINE sequences²⁴ are indicated by the letter L. The numbers beneath the schematic diagram give the designations for five of the unique-sequence subclones used to determine the extent of deletions in DNA samples from nine DMD males. After digestion with *Pst*I, each DNA sample was tested for the presence or absence of various human *Pst*I fragments by hybridization²² with the various subclones. In the two different autoradiographs presented, the hybridizing *Pst*I fragment for each subclone is indicated to the left of the figure; these are evident in lanes 2, 4 and 12, where the DNA samples isolated from individuals who do not bear deletions of DXS164 subclones are presented. The sizes of the respective *Pst*I hybridizing fragments for normal individuals are given to the right of the figure. The DXS164 subclones pERT87-18, -15 and -27 were radiolabelled separately¹¹ and hybridized simultaneously to a nitrocellulose filter containing the 12 immobilized *Pst*I-cleaved DNA samples. The DXS164 subclones pERT87-7 and -23 were also radiolabelled separately and hybridized together to a duplicate nitrocellulose membrane with the same immobilized *Pst*I-cleaved DNA samples. The different hybridization intensities of the *Pst*I fragments were due to differences in specific activity and hybridization characteristics of the individual probes. Lines 4 and 12 contained 5 µg of *Pst*I-cleaved DNA, whereas the others each contained ~1 µg. Lanes 1, 6 and 8-10 show no hybridization for any of the five DXS164 subclones. Lanes 3, 5, 7 and 11 (arrowed) exhibit at least one DXS164-hybridizing *Pst*I fragment. The deletion DNA sample in lane 3 exhibits an absence of the subclones (left-hand side), pERT87-27, -15 and -18 and has a single altered *Pst*I fragment for one of the subclones on the right-hand side (either pERT87-7 or -23). The altered fragment is presumably detected by the 23 subclone, and the hybridization results indicate that the deletion breakpoint occurs near pERT87-23 and the deletion extends in a leftward direction. The DNA sample in lane 7 exhibits similar hybridization results to the sample in lane 3, but the pERT87-23 hybridizing *Pst*I fragment is normal in size. This DNA sample is presumed to have a different deletion breakpoint between pERT87-7 and pERT87-23 compared with that in lane 3 and the deletion also extends to the left. The DNA sample in lane 5 exhibits a normal-sized *Pst*I hybridizing fragment for pERT87-23, a rare allele of a *Pst*I RFLP detected by pERT87-7 (~10% of both normal and DMD males exhibit this larger *Pst*I fragment) and normal-sized hybridizing *Pst*I fragments for pERT87-18 and -15. The pERT87-27 subclone exhibited no hybridization. This DMD male has not been previously shown to have a deletion of pERT87-8 or -1 and has a deletion which starts between pERT87-15 and pERT87-27 and extends also to the left. The DNA sample in lane 11 carries a deletion which breaks near that presented in lane 5, but the deletion extends in the opposite direction. pERT87-27 hybridizes to a normal-sized *Pst*I fragment whereas the remaining subclones are absent.

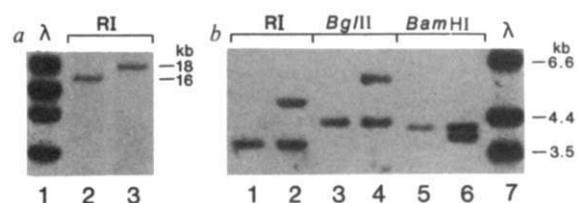


Fig. 3 Demonstration of deletion breakpoints for two independent deletions. Restriction enzyme cleavage, radiolabelling and hybridization were as described in Fig. 2 legend and ref. 1. In *a*, a deletion sample previously shown to carry pERT87-1 and a normal-sized hybridizing *Pst*I fragment, showed no hybridization for pERT87-8 or other subclones to the left of this one. Subclone pERT87-14, between pERT87-1 and -8, was used to detect by hybridization an altered *Eco*RI fragment in the deletion DNA sample (lane 2; compare with a normal-sized *Eco*RI fragment in lane 3). The normal *Eco*RI fragment stretches across pERT87-8 and, based on other restriction enzyme sites in the region, the breakpoint of this deletion has been determined to fall between the right side of pERT87-8 and a *Pst*I site 1.5 kb farther to the right. *b*, Demonstration of a deletion breakpoint detected with the pERT87-1 subclone. The deletion sample originally shown to be a deletion for pERT87-15 and -8 exhibits a break just to the left of pERT87-1. Lanes 1, 3 and 5, *Eco*RI, *Bgl*II and *Bam*HI digests, respectively, of a normal male; lanes 2, 4 and 6, *Eco*RI, *Bgl*II and *Bam*HI digests, respectively, of DNA isolated from a female heterozygous for the deletion. Both *Eco*RI and *Bgl*II reveal a larger altered fragment than *Bam*HI. The deletion breakpoint lies ~500-1,000 bp to the left of pERT87-1.

Figure 3 shows two such determinations for two independent deletions. In one case (Fig. 3*a*, lane 2), an altered hybridizing *Eco*RI fragment is detected with the pERT87-14 subclone (located between pERT87-8 and pERT87-1). Comparison of the size of the altered *Eco*RI fragment with that seen in normal individuals (Fig. 3*a*, lane 3) and the knowledge that pERT87-8 is absent indicate that this deletion breakpoint maps between the pERT87-8 segment and an unaltered *Pst*I site 1.5 kb towards pERT87-14. In the other example, different-sized hybridizing fragments for *Eco*RI (5.0 kb), *Bgl*II (5.8 kb) and *Bam*HI (3.9 kb) were observed in the DNA sample isolated from a female heterozygous for a deleted X chromosome. A sample isolated from an unaffected male (Fig. 3*b*, lanes 1, 3, 5) exhibited the normal hybridizing fragments for the three enzymes.

Most deletions were found to be larger than 137 kb, for no subclone from within the entire cloned region was found to be present. Among the 57 deletions tested, 24 showed the presence of some subclones but not others (the data are summarized in Fig. 4); these were deletions which had breakpoints within previously cloned DNA of the DXS164 locus. Fourteen deletions break on the left side of the cloned DNA and extend in the rightward direction off the DXS164 map towards the centromere (the direction of the DXS164 map relative to the centromere was established by analysing some deletion samples with the XJ-1.1 cloned segment; see Fig. 4 legend). Nine deletions break on the right of the map and extend in a leftward direction away from the centromere. Although some deletions are presented as similar, most have been shown to have independent breakpoints within the DXS164 locus (see Fig. 4 legend) and are schematically presented together when the breakpoints are near each other. It is apparent from Fig. 4 that there is a common region which is absent from almost all deletions except the six shown in the lower part of the figure. This region of common deletion overlap may be slightly biased due to the fact that the first probe made available to other investigators was pERT87-8, which resides near the overlap region. Further studies on additional DMD boys using all the DXS164 subclones should confirm whether this common region of overlap contains DNA sequences important for the expression of the DMD phenotype, or whether it represents an ascertainment bias.

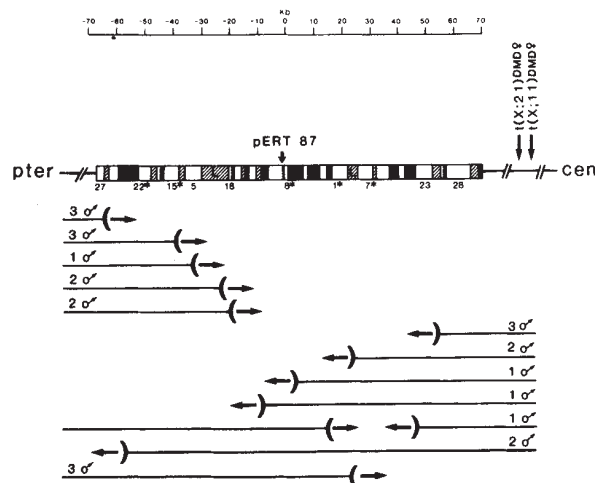


Fig. 4 A schematic representation of deletion breakpoints in the DXS164 locus. Fifty-seven DNA samples (53 samples shown to be deletions of one of the distributed subclones and 4 additional DNA samples shown to be deletions of either XJ-1.1 (ref. 2) or other DXS164 subclones) were tested for the absence or presence of various DXS164 subclones as presented in Figs 2, 3. Twenty-four of the 57 DNA samples showed the presence of at least one DXS164-hybridizing *Pst*I fragment, whereas the remaining 33 DNA samples exhibited no DXS164-hybridizing *Pst*I fragment—the entire 137 kb was missing. By using a combination of subclones including those exemplified in Fig. 2, the approximate positions of the 24 different breaks were determined and are indicated by a bracket with an arrow indicating the direction of the missing DNA. Certain breakpoints which were similar are combined, and the number of males combined is indicated at the side of the figure. Nine deletion breakpoints were determined to the extent where an aberrant hybridizing restriction fragment was identified; these are localized within a few kilobases within the DXS164 locus. Nine deletions were determined to be different from the first nine, as revealed by the pattern of their hybridization with DXS164 subclones, but an altered restriction fragment has yet to be identified and these are thus only localized to within 1–5 kb. Six deletion breakpoints (the most recently acquired) were only mapped to 10–30-kb regions of the DXS164 locus and have not been definitely shown to differ from the other 18. Some of the deletions presented have also been tested for hybridization with the previously described clone XJ-1.1 (ref. 2): of those tested that break within DXS164 and extend to the right, all show no hybridization with XJ-1.1. For those tested that break in DXS164 and extend to the left, XJ-1.1 is present. The hybridization results with XJ-1.1 allow localization of the DXS164 subclones relative to the centromere as well as the (X; 21) translocation breakpoint, as indicated in the figure. Four of the six deletion DNA samples depicted at the bottom of the figure were not determined to be deletions by the original subclones pERT87-1, -8 or -15; two were found to be deletions for the XJ-1.1 cloned segment and were subsequently mapped with additional DXS164 clones, and the remaining two deletions were detected only with the leftmost DXS164 subclone, pERT87-27.

Screening of boys with additional pERT87 subclones and with the XJ-1.1 clone detected four deletions which would not have been detected with the originally distributed clones. These deletions were found among 269 boys who had been shown previously not to carry deletions of the pERT87-1, -8 or -15 subclones. These four deletions are not included in the data presented in Table 1 but are shown in Fig. 4; it is assumed that the true proportion of DMD and BMD boys who have deletions of portions of their X chromosomes would be greater than that presented in Table 1, where only three probes were used. Analysis of DMD or BMD DNA samples with the most distal pERT87-27 subclone, the three central subclones pERT87-1, -8 and -15, and the XJ-1.1 subclone should detect most of the deletions that are currently detectable.

The breakpoints of six deletions are different from most others; these are shown at the bottom of Fig. 4. One deletion is completely encompassed within cloned DNA and removes only 45 kb of DNA. Three deletions break within this 45-kb region and extend off the map towards the centromere; two of these were originally detected by the more centromeric cloned segment XJ-1.1 and subsequently mapped within the DXS164 locus. Two other deletions (detected with subclones of DXS164 other than pERT87-1, -8 and -15) break at the extreme left end of the cloned DNA and extend off the map for an undetermined distance. The central location of DXS164 subclones 1, 8 and 15 to DMD and/or BMD mutational sites, inferred from family studies and translocation breakpoint mapping, is substantiated by these additional deletion DNA samples. The profile of deletion breaks indicates that there must be a large segment of DNA which, when disrupted, can yield the phenotype of DMD and/or BMD. The fact that large deletions yield a DMD phenotype similar to that of other DMD boys who have not been demonstrated to bear deletions, indicates that the product of the locus can be either completely absent or aberrant and still yield a similar clinical picture.

The pattern of deletions implies that the gene(s) responsible for DMD must lie on a very large segment of X-chromosomal DNA. This is substantiated by the fact that recombinants are observed between cloned segments in the region and some disease mutation sites. For the future of diagnosis, the deletions represent a ready means of obtaining additional cloned segments from the other side of those deletions that break within DNA that has already been cloned. These junctional fragments from deletion patients are presently being cloned and the development of new RFLP-detecting probes to mark the entire region in which mutation can yield DMD and BMD is imminent. The deletion breakpoints also demarcate the regions of the DXS164 locus that might contain information important for the expression of the gene responsible for DMD and/or BMD. Analysis of these breakpoints at the level of the nucleotide sequence may also elucidate the mechanism underlying the deletions.

We thank all the families and patients involved in this study. We also acknowledge the many funding agencies who provided support for this work, especially the Muscular Dystrophy Association of America, the Muscular Dystrophy Group of Great Britain and Northern Ireland, the Muscular Dystrophy Association of Canada, and members of the European Alliance of Muscular Dystrophy Associations. We also thank the clinical geneticists, paediatricians and neurologists who brought to our attention the many Duchenne and Becker muscular dystrophy patients.

Received 26 March; accepted 1 May 1986.

1. Monaco, A. P. *et al.* *Nature* **316**, 842–845 (1985).
2. Ray, P. N. *et al.* *Nature* **318**, 672–675 (1985).
3. Bender, W. *et al.* *Science* **221**, 23–29 (1983).
4. Frischauf, A. M. *et al.* *J. molec. Biol.* **170**, 827–842 (1983).
5. Botstein, D. *et al.* *Am. J. hum. Genet.* **32**, 314–331 (1980).
6. Brown, C. S. *et al.* *Hum. Genet.* **71**, 62–74 (1985).
7. Fadda, S. *et al.* *Hum. Genet.* **71**, 33–36 (1985).
8. Wilcox, D. E. *et al.* *Hum. Genet.* **70**, 365–378 (1985).
9. de la Chapelle, A. (ed.) *Human Gene Mapping 8* (Karger, Basel, 1985).
10. Hofker, M. H. *et al.* *Hum. Genet.* **70**, 148–156 (1985).
11. Aldridge, J. *et al.* *Am. J. hum. Genet.* **36**, 546–564 (1984).
12. de Martinville, B. *et al.* *Am. J. hum. Genet.* **37**, 235–249 (1985).
13. Francke, U. *et al.* *Am. J. hum. Genet.* **37**, 250–267 (1985).
14. Kunkel, L. M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **82**, 4778–4782 (1985).
15. Drayna, D. *et al.* *Science* **230**, 753–758 (1985).
16. Mohandas, T. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **77**, 6759–6763 (1980).
17. Worton, R. *et al.* *Science* **224**, 1447–1449 (1984).
18. Jacobs, P. A. *et al.* *Am. J. hum. Genet.* **33**, 531–538 (1981).
19. Pearce, J. M. S. *et al.* *J. Neurol. Neurosurg. Psychiatr.* **27**, 181–185 (1964).
20. Emery, A. E. H. & Holloway, S. *Hum. Hered.* **27**, 118–126 (1977).
21. Appel, S. H. & Roses, A. D. in *The Metabolic Basis of Inherited Disease* (ed. Stanbury, J. B.) 1470–1495 (McGraw-Hill, New York, 1983).
22. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
23. Middlesworth, W. *et al.* *Nucleic Acids Res.* **13**, 5723 (1985).
24. Singer, M. F. *Int. Rev. Cytol.* **76**, 67–112 (1982).

Carcinogen-specific mutation and amplification of Ha-ras during mouse skin carcinogenesis

Miguel Quintanilla, Ken Brown,
Martin Ramsden* & Allan Balmain

Beatson Institute for Cancer Research, Switchback Road,
Bearsden, Glasgow G61 1BD, UK

Cellular proto-oncogenes can be activated by both point mutations and chromosomal translocations, suggesting that there may be a direct link between exposure to agents which damage DNA and genetic change leading to malignancy (see refs 1, 2 for reviews). Several groups have therefore analysed mutations found in cellular oncogenes of tumours induced by particular physical or chemical carcinogens³⁻⁹. Here, we have analysed the molecular changes at different stages of carcinogenesis in mouse skin tumours induced by initiating and promoting agents. Over 90% of tumours, including premalignant papillomas, initiated with dimethylbenzanthracene (DMBA) have a specific A→T transversion at the second nucleotide of codon 61 of the Harvey-ras (Ha-ras) gene. The frequency of this mutation was dependent on the initiating agent used, but not on the promoter, suggesting that the mutation occurs at the time of initiation. The mutation was heterozygous in most papillomas tested, but was homozygous or amplified in some carcinomas. The development of further chromosomal changes at the c-Ha-ras gene locus is therefore a common feature of tumour progression.

One of the consequences of *ras* gene mutation is that the resultant altered proteins (p21) exhibit aberrant mobilities on SDS gels^{10,11}. Alterations at codon 12 give rise to more slowly migrating forms of p21, while mutations at the other frequently affected locus around codon 61 usually produce rapidly migrating forms. Figure 1B (lane 5) shows the normal Kirsten-ras (Ki-ras) p21 immunoprecipitated from NIH 3T3 cells by monoclonal antibody Y13-259 (ref. 12). Transformants induced by transfection with DNA from a papilloma (Fig. 1B, lane 1) or a carcinoma (lane 3) synthesize in addition a p21 product encoded by the transforming Ha-ras gene, as shown using the Harvey-specific monoclonal antibody YA6-172, which immunoprecipitates only the faster migrating species (Fig. 1A, lane 1). Although the parental NIH 3T3 cells produce only small amounts of Harvey p21, a band can be detected after immunoprecipitation with YA6-172, which migrates at the front edge of the Ki-ras p21 band (Fig. 1A, lane 3). Comparison of lanes 1 and 3 of Fig. 1A indicates that the transforming Ha-ras encoded by the tumour DNAs has a substantially higher mobility on SDS gels and probably results from a mutation in the region of the gene surrounding codon 61. In an analysis of representative foci induced by seven tumours, six were shown to synthesize rapidly migrating forms of Ha-ras p21; four of these are shown in Fig. 1C, together with a slowly migrating p21 (lane 5) in a transformant induced by DNA from a Sencar carcinoma. We conclude from this analysis that most skin tumours initiated with DMBA probably have mutations around codon 61 of the mouse Ha-ras gene, but a minority exhibit mutations at or around codon 12.

Restriction fragment length polymorphisms (RFLPs) have been used to demonstrate directly the presence of mutated *ras* genes in samples of DNA from human or animal tumours^{7,13,14}. To attempt this experimental approach, we first obtained a genomic clone of the normal mouse cellular Ha-ras gene and sequenced coding exons 1 and 2 which contained, respectively, the putative mutable sites at codons 12 and 61 (M.R., K.B., F.

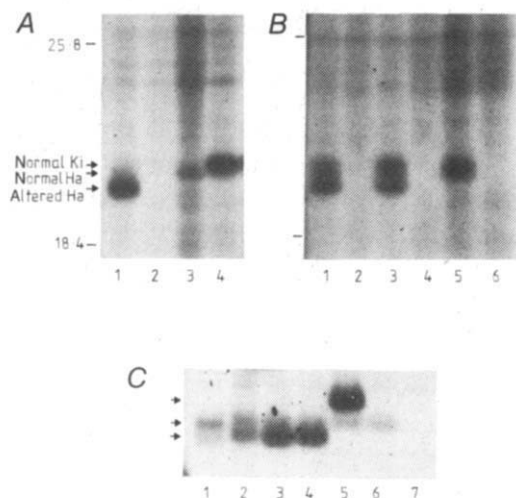


Fig. 1 Immunoprecipitation of *ras* p21 from NIH 3T3 transformants. **A**, Lanes 1, 2: NIH 3T3 transformant obtained by transfection with DNA from a primary carcinoma. Lanes 3, 4: NIH 3T3 cells. Lysates were immunoprecipitated with antibody YA6-172 (lanes 1, 3), normal rat serum (lane 2) or antibody Y13-259 (lane 4). **B**, NIH 3T3 transformants obtained by transfection with DNA from a papilloma (lanes 1, 2) or a carcinoma (lanes 3, 4). Lanes 5, 6: normal NIH 3T3 cells. Lysates in lanes 1, 3 and 5 were immunoprecipitated with Y13-259 and those in lanes 2, 4 and 6 with normal rat serum. **C**, NIH 3T3 transformants obtained by transfection with DNA from individual papillomas (lanes 1-3) or carcinomas (lanes 4, 5), all immunoprecipitated with YA6-172. Lanes 6 and 7: lysates from NIH 3T3 cells immunoprecipitated with YA6-172 or normal rat serum respectively.

Methods. Isolation of DNA by lysis of frozen tumours in 5 M guanidinium thiocyanate and centrifugation through CsCl has been described previously^{3,9}. NIH 3T3 transformants obtained by calcium phosphate-mediated transfection of DNA from primary tumours were grown to subconfluence in Special Liquid Medium (Flow) containing 10% fetal bovine serum. Cultures were labelled by incubation for 18 h in methionine-free medium containing 5% dialysed fetal bovine serum in the presence of ³⁵S-methionine (NEN; specific activity 1,000 Ci mmol⁻¹, 200 µCi per ml of medium). Cells were lysed essentially as described by Furth *et al.*¹² and aliquots (3–5 × 10⁷ acid-precipitable c.p.m.) of the lysate incubated with monoclonal antibody YA6-172 or Y13-259, or normal rat serum. Antigen-antibody complexes were precipitated with protein A-Sepharose coated with rabbit anti-rat IgG (Sera Labs). Precipitates were washed six times in lysis buffer and run on 12.5% polyacrylamide gels as described by Laemmli²⁵. Gels were treated for fluorography with Enlightening (NEN), dried and exposed to Kodak film at –70 °C using intensifying screens.

Fee, R. Krumlauf and A.B., manuscript in preparation). The sequence around codon 61, which appeared to be the most frequently mutated locus in DMBA-initiated tumours, is shown in Fig. 2B. Clearly, an A→T transversion at this position could give rise to a new restriction site for the enzyme *Xba*I. Digestion with *Xba*I of DNA samples from the NIH 3T3 transformants shown in Fig. 1 revealed the presence of novel hybridizing fragments in all but one of the foci tested (Fig. 2A, lane a). This particular focus is that which exhibited the slowly migrating p21 (Fig. 1C), presumably because of a codon 12 mutation, and would therefore not be expected to show any change in the sequence of codon 61. Most foci had novel fragments of ~4 and 8 kilobases (kb) but some had aberrant bands due to the loss of one or more of the flanking *Xba*I sites during integration of the transfected DNA (Fig. 2A, lanes f–h).

The *Xba*I polymorphism enabled us to verify that the codon 61 mutation was present in the primary mouse skin tumours, and was not an artefact of the transfection procedure. Table 1 shows that 90% of DMBA-initiated tumours from three mouse strains had the same mutation. Figure 3A, lanes a and e, shows

* Present address: Glaxo Research Laboratories, Ulverston, Cumbria LA12 9DR, UK



Fig. 2 *Xba*I polymorphism in NIH 3T3 transformants. **A**, DNAs derived from NIH 3T3 transformants obtained by transfection with papilloma (lanes *d-h,j*) or carcinoma (lanes *a-c*) DNAs were digested with *Xba*I and hybridized on Southern blots with an *Ha-ras*-specific probe. Lane *a* contains DNA from a transformant with the slowly migrating p21 protein shown in Fig. 1C. Lane *i* contains DNA from a spontaneous NIH 3T3 transformant and has the same hybridization pattern as normal DNA. **B**, Induction of an *Xba*I polymorphism by a specific mutation at codon 61. Digestion of normal mouse DNA with *Xba*I gives rise to a 12-kb fragment containing the mouse *c-Ha-ras* gene. Mutation of the middle base of codon 61 from A to T would give rise to a new *Xba*I site, generating tumour-specific fragments of ~8 and 4 kb.

Methods. Aliquots (10 µg) of DNA from different transformants were digested with a 5–10-fold excess of *Xba*I (BioRad) under the conditions recommended by the suppliers. DNAs were run on 1% agarose gels, and transferred to nitrocellulose filters (Sartorius) according to Law *et al.*²⁶. Purified insert from the plasmid BS9 (ref. 27) was nick-translated and blots were hybridized at 42 °C in 50% formamide buffer as described previously³. Blots were normally washed to a final stringency of 0.5×SSC at 65 °C, then exposed to Kodak X-OMAT film for 1–3 days at –70 °C with intensifying screens.

the typical pattern observed in 90% of the papillomas. Both the normal 12-kb band and those derived from the mutated allele are evident, and in almost all cases the band corresponding to the normal fragment is the most intense. Since histological analysis shows that papillomas consist predominantly of epithelial cells with a relatively small proportion of contaminating stroma cells, we conclude that the mutation is probably heterozygous in these premalignant tumours. A small proportion (~10%) of the papillomas tested showed no evidence of the *Xba*I polymorphism (see, for example, Fig. 3A, lane *d*).

Figure 3A also shows the results of an analysis of a series of carcinomas initiated with DMBA. All the carcinomas from NIH mice exhibited the *Xba*I polymorphism, but in some cases, the relative intensity of the bands diagnostic of the mutated allele was considerably stronger than in the papillomas (Fig. 3A, lanes *b,f-h,k*). Figure 3B shows the pattern obtained after rehybridization of the same blot with a single-copy probe for the mouse interleukin-3 (IL-3) gene¹⁵. Comparison of the relative band intensities demonstrates that the mutated allele of the *c-Ha-ras* gene has become amplified in the carcinomas shown in lanes *f-h*. Those shown in lanes *b* and *k*, which contain more DNA and therefore do not provide clear evidence of amplification, nevertheless exhibit a change in the relative amounts of 'normal' 12-kb and 'mutated' 4- and 8-kb fragments. This result suggests that homozygosity may have developed in these cases, although

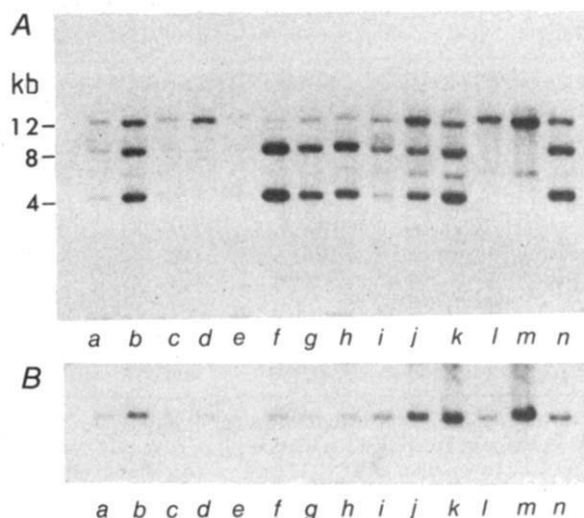


Fig. 3 Demonstration of the *Xba*I polymorphism in primary tumours. **A**, Lanes *a, d* and *e*: DNAs obtained from DMBA-initiated primary papillomas. Lanes *b, c* and *f-k* contain NIH carcinoma DNAs. The pattern obtained with normal DNA is shown in lane *l*. Lanes *m* and *n* contain respectively DNA from spontaneously transformed NIH 3T3 cells and an NIH 3T3 transformant obtained by transfection with DNA from a papilloma with the codon 61 polymorphism. DNAs were digested with *Xba*I, run on agarose gels and hybridized with an *Ha-ras*-specific probe as described in Fig. 2 legend. **B**, Autoradiograph showing the ~1-kb band obtained by rehybridization of the blot shown in **A** with a complementary DNA probe for the IL-3 gene¹⁵ (provided by A. Dunn).

loss of the normal allele is difficult to prove because of the presence of some normal cells in the tumours. Yet another pattern is seen in the carcinomas in lanes *c, i* and *j*, which have similar band ratios to the papillomas.

The high frequency of the *Xba*I polymorphism in DMBA-initiated tumours allowed us to determine whether this mutation is carcinogen-specific, or is detected in tumours induced by different carcinogenic initiators or promoters. Chryserobin is a strong tumour promoter and a derivative of anthralin¹⁶. Compounds of this type are structurally unrelated to tetradecanoyl phorbol 13-acetate (TPA), do not bind to the phorbol ester receptor¹⁷ and presumably have a different site of action within the target cell. The results shown in Table 1 suggest that the nature of the promoter used does not influence the frequency of the specific *Xba*I polymorphism seen in DMBA-initiated tumours. However, of a total of 12 tumours from NIH mice initiated with *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG), comprising six papillomas and six carcinomas, none showed the restriction fragments diagnostic of the mutation. The carcinogen MNNG is a direct-acting methylating agent and might therefore not be expected to induce the same types of mutation as DMBA¹⁸. Since the mice were initiated by a single treatment with the carcinogen, the presence of the polymorphism in tumours initiated with one type of initiator, but not with another, suggests that the *Ha-ras* gene mutations occur at or very close to the time of initiation.

Zarbl *et al.*⁷ recently showed that specific G→A transitions occur consistently at the second nucleotide of codon 12 of the rat *c-Ha-ras* gene in rat mammary carcinomas induced by a single treatment with nitrosomethylurea (NMU). As this particular mutation is exactly that predicted on the basis of the known mechanism of action of NMU¹⁸, Zarbl *et al.*⁷ concluded that NMU interacts directly with codon 12 of the *ras* gene. Our observation that *Ha-ras* mutations in DMBA-initiated tumours occur predominantly at adenosine residues is compatible with previous results on the metabolism and binding of DMBA to

Table 1 *Xba*I polymorphism in Ha-ras gene in skin tumours

Strain	Tumour	Initiator	Promoter	No. positive no. tested
Sencar	Papilloma	DMBA	TPA	12/14
		DMBA	Chryserobin	5/5
NMRI	Papilloma	DMBA	TPA	2/2
NIH	Papilloma	DMBA	TPA	3/3
Sencar	Carcinoma	DMBA	TPA	0/2*
NMRI	Carcinoma (transplanted)	DMBA	TPA	3/3
NIH	Carcinoma	DMBA	TPA	8/8
NIH	Papilloma	MNNG	TPA	0/6
NIH	Carcinoma	MNNG	TPA	0/6

Mice were initiated with 25 µg DMBA or 600 µg MNNG, dissolved in 200 µl acetone. Promotion was carried out twice weekly by treatment with an acetone solution of TPA (12.5 µg) or chryserobin (50 µg)¹⁶.

* One tumour has an Ha-ras mutation, presumably at codon 12 (see Fig. 1C, lane 5).

DNA in mouse skin. Dipple and co-workers have shown that benzo(a)pyrene, which is about 30-fold weaker as an initiating agent than DMBA, exhibits a corresponding decrease in the amount of binding to deoxyadenosine, but not deoxyguanosine, residues^{19,20}. The relative carcinogenic potency of these two aromatic hydrocarbons may be related to their abilities to bind to, and therefore induce mutations at, deoxyadenosine. The mutations in DMBA-induced tumours are not exclusively A → T transversions, however, as a small proportion of tumours are activated at codon 12, presumably by changes at one of two G residues.

If Ha-ras gene mutation is critically involved in initiation of mouse skin carcinogenesis, the introduction of a mutated gene into epidermal cells *in vivo* should be able to substitute for treatment with the initiating chemical. We have shown recently that the viral Ha-ras gene of Harvey sarcoma virus can act as an initiator of two-stage skin carcinogenesis²¹. These results support the conclusion that Ha-ras gene mutation occurs during initiation, but this event in itself is insufficient for tumorigenesis because even epidermal cells initiated with the viral Ha-ras gene required promoter treatment before papillomas developed²¹.

These studies suggest that mutation and amplification of *ras* genes may occur at different stages of carcinogenesis. It is possible that the additional changes involving the mutated allele may take place at the time of carcinoma development, or may be present in a sub-set of papillomas with a high probability of malignant conversion²². Indeed, one of 22 papillomas in the present studies showed amplification of the mutated allele (data not shown). We do not know whether these changes are accompanied by loss of the normal allele of the Ha-ras gene; loss of the normal allele has been seen in a carcinogen-induced thymic lymphoma²³ and in human tumour cell lines¹³, and loss of one Ha-ras allele has been reported in a series of bladder carcinomas²⁴. Chromosomal changes of this kind may represent late events in tumour development in some, but by no means all, mouse skin carcinomas.

M.Q. is the recipient of an EMBO long-term Fellowship and K.B. is supported by a Training Fellowship from the MRC. We thank I. B. Kerr for histological diagnosis of tumours, Francis Fee and Sheena Young for technical assistance, and J. Paul for critical reading of the manuscript. Some of the tumours used in this study were supplied by G. T. Bowden, T. Slaga and J. Di Giovanni. The Beatson Institute is funded by the Cancer Research Campaign.

Received 26 February; accepted 7 May 1986.

- Klein, G. & Klein, E. *Carcinogenesis* **5**, 429-435 (1984).
- Balmain, A. *Br. J. Cancer* **51**, 1-7 (1985).
- Balmain, A. & Pragnell, I. B. *Nature* **303**, 72-74 (1983).

- Sukumari, S., Notario, V., Martin-Zanca, D. & Barbacid, M. *Nature* **306**, 658-661 (1983).
- Guerrero, I., Calzada, P., Mayer, A. & Pellicer, A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 202-205 (1984).
- Eva, A. & Aaronson, S. *Science* **220**, 955-956 (1983).
- Zarbl, H., Sukumari, S., Arthur, A. V., Martin-Zanca, D. & Barbacid, M. *Nature* **315**, 382-386 (1985).
- Guerrero, I., Villasante, A., Corcos, V. & Pellicer, A. *Science* **225**, 1159-1162 (1984).
- Balmain, A., Ramsden, M., Bowden, G. T. & Smith, J. *Nature* **307**, 658-660 (1984).
- Seeburg, P. H., Colby, W. W., Capon, D. J., Goeddel, D. V. & Levinson, A. D. *Nature* **312**, 71-75 (1984).
- Yuasa, Y. *et al. Nature* **303**, 775-779 (1983).
- Furth, M. E., Davis, L. J., Fleudelys, B. & Scolnick, E. M. *J. Virol.* **43**, 294-304 (1982).
- Santos, E. *et al. Science* **223**, 661-664 (1984).
- Fujita, J. *et al. Nature* **307**, 464-466 (1984).
- Fung, M. C. *et al. Nature* **307**, 233-237 (1984).
- DiGiovanni, J. & Boutwell, R. K. *Carcinogenesis* **4**, 281-284 (1983).
- Delclos, K. B., Nagle, D. S. & Blumberg, P. M. *Cell* **19**, 1025-1032 (1980).
- Margison, G. P. & O'Connor, P. J. *Chemical Carcinogens and DNA* Vol. 1 (ed. Grover, P.) (CRC Press, Florida, 1978).
- Sawicki, J. T., Moschel, R. C. & Dipple, A. *Cancer Res.* **40**, 1580-1582 (1983).
- Dipple, A., Sawicki, J. T., Moschel, R. C. & Bigger, C. A. H. in *Extrahepatic Drug Metabolism and Chemical Carcinogenesis* (eds Rystrom, J., Montelius, J. & Bengtsson, M.) 439-448 (Elsevier, Amsterdam, 1983).
- Brown, K. *et al. Cell* (in the press).
- Hennings, H., Shores, R., Mitchell, P., Spangler, E. F. & Yuspa, S. H. *Carcinogenesis* **6**, 1607-1610 (1985).
- Guerrero, I., Villasante, A., Corcos, V. & Pellicer, A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7810-7814 (1985).
- Fearon, E. R., Feinberg, A. P., Hamilton, S. H. & Vogelstein, B. *Nature* **318**, 377-380 (1985).
- Laemmli, U. K. *Nature* **227**, 680-682 (1970).
- Law, D. J., Frossard, P. M. & Rucknagel, D. L. *Gene* **28**, 153-158 (1984).
- Ellis, R. W. *et al. J. Virol.* **36**, 408-420 (1980).

Site-directed mutagenesis of the regulatory light-chain $\text{Ca}^{2+}/\text{Mg}^{2+}$ binding site and its role in hybrid myosins

Fernando C. Reinach*, Kiyoshi Nagai† & John Kendrick-Jones†

* Ludwig Institute for Cancer Research and † MRC Laboratory of Molecular Biology, MRC Centre, Hills Road, Cambridge CB2 2QH, UK

The regulatory light chains, small polypeptides located on the myosin head, regulate the interaction of myosin with actin in response to either Ca^{2+} or phosphorylation. The demonstration that the regulatory light chains on scallop myosin can be replaced by light chains from other myosins has allowed us to compare the functional capabilities of different light chains¹, but has not enabled us to probe the role of features, such as the $\text{Ca}^{2+}/\text{Mg}^{2+}$ binding site, that are common to all of them. Here, we describe the use of site-directed mutagenesis to study the function of that site. We synthesized the chicken skeletal myosin light chain in *Escherichia coli* and constructed mutants with substitutions within the $\text{Ca}^{2+}/\text{Mg}^{2+}$ binding site. When the aspartate residues at the first and sixth Ca^{2+} coordination positions are replaced by uncharged alanines, the light chains have a reduced Ca^{2+} binding capacity but still bind to scallop myosin with high affinity. Unlike the wild-type skeletal light chain which inhibits myosin interaction with actin, the mutants activate it. Thus, an intact $\text{Ca}^{2+}/\text{Mg}^{2+}$ binding site in the N-terminal region of the light chain is essential for regulating the interaction of myosin with actin.

The myosin regulatory light chains belong to a family of calcium-binding proteins which includes calmodulin and troponin C²⁻⁴. These are composed of four EF hands or domains, each consisting of an α -helix E, a divalent cation binding loop and an α -helix F (refs 5-7). In the myosin regulatory light chains, deletions and non-conservative substitutions at critical positions have destroyed the $\text{Ca}^{2+}/\text{Mg}^{2+}$ binding ability of three of the loops, so that only the N-terminal domain has retained the ability to bind divalent ions⁸⁻¹¹.

The regulatory light chains are non-covalently bound to the heavy chains in the neck region of myosin heads¹²⁻¹⁵. They can be subdivided into three functional classes: (1) in molluscan

muscles, they inhibit myosin interaction with actin in the absence of Ca^{2+} and Ca^{2+} relieves this inhibition¹; (2) in vertebrate smooth muscle and non-muscle cells, Ca^{2+} binding to calmodulin activates a kinase which phosphorylates the regulatory light chains and thus 'switches on' myosin interaction with actin^{16,17}; (3) in vertebrate striated muscles, they do not appear to have a primary role in actomyosin regulation^{18,19}.

When the regulatory light chains are removed from scallop myosin by EDTA, myosin interaction with actin as measured by the actin-activated myosin MgATPase is 'desensitized' to Ca^{2+} and Ca^{2+} sensitivity can be restored when the scallop light chains rebind^{20,21} (Table 1). Regulatory light chains from a wide

Table 1 Effect of the light-chain mutants on Ca^{2+} regulation in desensitized scallop myofibril-regulatory light chain hybrids

Desensitized scallop myofibrils	Actomyosin MgATPase	
	- Ca^{2+}	+ Ca^{2+}
	($\mu\text{mol H}^+ \text{min}^{-1} \text{mg}^{-1}$)	
Control	0.163	0.152
+ scallop LC	0.050	0.263
+ rabbit LC	0.050	0.081
+ chicken LC	0.058	0.102
+ <i>E. coli</i> LC	0.050	0.069
+ mutant 442 LC	0.262	0.233
+ mutant 452 LC	0.229	0.200
+ mutant 463 LC	0.251	0.231

The desensitized scallop myofibril-regulatory light chain hybrids were prepared as described in Fig. 2B legend. Their actin-activated MgATPase activities were measured at 25 °C in a final concentration of 36 mM KCl, 3 mM MgCl_2 , 2 mM MgATP and either 0.1 mM EGTA ($-\text{Ca}^{2+}$) or 0.1 mM EGTA and 0.15 mM Ca^{2+} ($+\text{Ca}^{2+}$) by the pH-stat assay procedure as described previously¹. Samples (1 mg) of each of the hybrids were solubilized in 0.6 M KCl, 1 mM MgATP, pre-mixed with 0.5 mg skeletal muscle F-actin (actin filaments) in a final volume of 200 μl , and then diluted to 5 ml to give the final assay conditions shown above. The mutant light chains remain bound to the desensitized scallop myosin under these assay conditions. These ATPase assays were repeated with three other desensitized scallop myofibril preparations and with the light chain/mutant hybrids prepared under the various conditions described in Fig. 2B legend. We observed that under all the conditions tested the native and wild-type vertebrate light chains always inhibited the actin-activated myosin ATPase whereas the mutant light chains always activated the ATPase. LC, light chain.

variety of myosins will also bind to desensitized scallop myosin to form hybrid myosins and their function can be assayed¹. The three classes of light chains regulate the actin-activated ATPase of hybrid scallop myofibrils as follows: those from molluscan muscles regulate in a calcium-dependent fashion, whereas those from vertebrate smooth muscle and non-muscle cells regulate in a phosphorylation-dependent manner¹⁷ and those from vertebrate striated muscles inhibit the ATPase activity²². There is no evidence that the divalent cation binding site in any of these three classes of regulatory light chains is involved in Ca^{2+} regulation^{11,23}, and thus the precise role of this site is not known. We have prepared chicken skeletal regulatory light-chain mutants with genetically engineered substitutions in the $\text{Ca}^{2+}/\text{Mg}^{2+}$ binding loop, determined the Ca^{2+} binding capacity of these mutants and assayed their regulatory function using desensitized scallop myofibrils.

A plasmid pLcIFXMLC was constructed which directed the synthesis of a fusion protein consisting of the N-terminal 31 amino-acid residues of λ cII protein, the blood coagulation factor X_a recognition site and the complete chicken myosin light-chain (MLC) polypeptide (Fig. 1). Because the amino-acid sequence near the factor X_a cleavage site, Arg-Ala-Pro-Lys, resembles the thrombin cleavage site in fibrinogen²⁴, we digested the fusion protein CIIIFXMLC with bovine thrombin. This enzyme cleaved the fusion protein at a single site to yield a

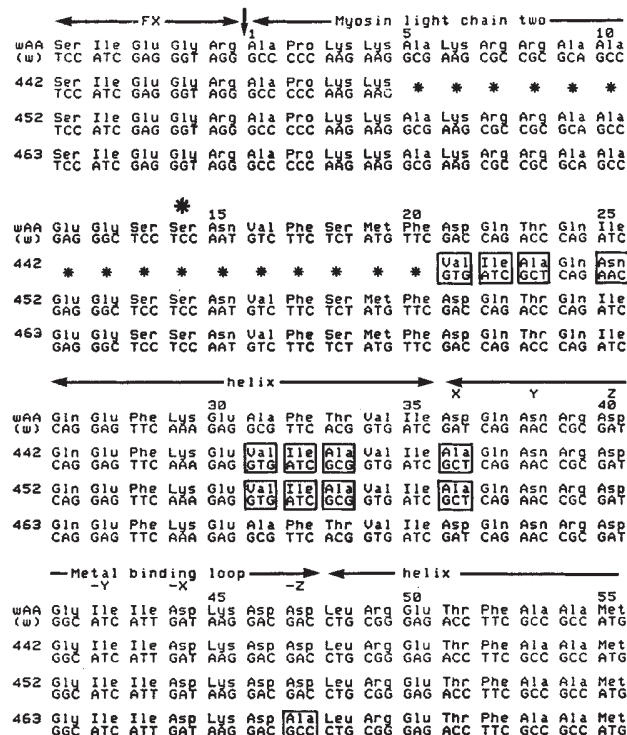


Fig. 1 Sequence of the factor X recognition site (FX) and the first 55 amino acids of the myosin regulatory light-chain clone present in pLcIFXMLC. Arrow, thrombin cleavage site. The nucleotide and amino-acid sequences of the wild-type light chain (wAA), mutants 442, 452 and 463 are indicated. Where the nucleotide and amino-acid sequences of the three mutant light chains are different from the wild type the residues are boxed. The deletion present in clone 442 is marked with asterisks. The residues involved in the α -helix-Ca/Mg binding loop- α -helix, which are believed to form an EF-hand structure, are indicated. The six residues believed to be involved in the Ca/Mg coordination are also indicated (X, Y, Z, -Y, -X, -Z). The phosphorylation site is indicated with a large asterisk.

Methods. Construction of the expression plasmid. The *EcoRI* insert of clone λ gt11/L10 containing the whole coding region of chicken myosin light chain 2 (see Fig. 5 of ref. 25 for sequence) was treated with the Klenow fragment of DNA polymerase in the presence of dNTP to obtain flush ends and cloned into the *SmaI* site of M13mp19. A clone (M13MLC) containing the insert with the 5' end of the messenger RNA facing the universal priming site, was selected. A *Sau96I* fragment of M13MLC from position 33 to 253 was isolated, treated with Klenow DNA polymerase in the presence of dNTP to obtain flush ends, and cloned in the *SmaI* site of M13mp11 (ref. 27). A clone (M13mp11FXMLC1) where the Ala codon of the insert was in-frame with the FX recognition sequence and facing the universal priming site of M13, was selected. The *NcoI* (position 194)-*HindIII* (polylinker) fragment from M13MLC was then cloned into the *NcoI*-*HindIII*-cleaved M13mp11FXMLC1 to obtain M13mp11FXMLC2. This clone contains the whole MLC coding and 3' noncoding region linked in-frame to the FX recognition site. The *BamHI* fragment from M13mp11 FXMLC2 containing FX sequences and MLC sequences up to position 364 was cloned into the *BamHI* site of pLcII (ref. 28) and a clone (pLcIFXMLC1) with the FX recognition sequence in-frame with the first 31 amino acids of cII was selected by restriction enzyme analysis. Finally, the *NcoI* (194)-*HindIII* (polylinker) fragment of M13MLC was cloned into *NcoI*-*HindIII*-cleaved pLcIFXMLC1 to obtain pLcIFXMLC2. This clone contains the pL promoter of λ followed by *nutL*, *nutR*, *trI*, and the first 31 amino acids of cII. The sequence at the junction of FX and MLC is shown. The *EcoRI*-*HindIII* fragment of pLcIFXMLC2 was subcloned into M13mp19 to obtain M13mp404. Single-stranded templates of M13mp404 were used to sequence the final construct with the help of synthetic oligonucleotides spaced along the insert. Construction of mutants. Single-stranded DNA from clone M13mp404 and oligonucleotides synthesized using the phosphotriester method (5'-GGTGATCGCTCAGAACCC to change Asp36 to Ala and 5'-TAAGGACGCCCTGCGGG also to change Asp47 to Ala) were hybridized and extended with Klenow DNA polymerase and transfected into BMH-71-18-*mutL* (ref. 29). Mutant clones were identified by hybridization with the mutagenic oligonucleotide²⁷, plaque-purified and the whole *EcoRI*-*HindIII* insert was sequenced before recloning into pLmp10 (ref. 28). Restriction enzymes were from New England Biolabs, Klenow DNA polymerase from Boehringer Mannheim and T4 DNA ligase was purified from strain NM989 (ref. 30). M13 constructs were grown in TG-1 (ref. 29) and pLcII constructs in QY13 (ref. 31).

Fig. 2 A, Purification of light-chain mutants; B, SDS-PAGE (10–20% acrylamide gradient gel) of desensitized scallop myofibril/regulatory light chain hybrids. A, SDS (15%)-PAGE of: (1) mutant 442; (2) mutant 452; and (3) mutant 463 at the different stages in the purification procedure: a, cell/pellet fractions after 4 washes with Triton X-100 solution; b, mutant light-chain fusion proteins cIIFX-MLC after purification on DEAE-cellulose; c, mutant light chains after removal of the bacterial leader sequence by thrombin cleavage and DEAE-cellulose chromatography. The gels were calibrated using the following M_r standards: 68,000 (68K), bovine serum albumin; 42K, actin; 25K, rabbit skeletal muscle myosin A₁ (LC₁) light chain; 18K, rabbit myosin regulatory (LC₂) light chain. B, The hybrids were as follows: 1, control, desensitized scallop myofibrils alone; 2, +scallop light chains; 3, +rabbit skeletal muscle myosin light chains; 4, +chicken breast muscle myosin light chains; 5, +*E. coli*-synthesized light chains (wild type); 6, +light-chain mutant 442; 7, +mutant 452; and 8, +mutant 463. The following M_r standards were used: 200K, myosin heavy chain; 94K, phosphorylase; 68K, bovine serum albumin; 42K, actin; 25K, rabbit myosin A₁ (LC₁) light chain; 19K, chicken myosin LC₂ light chain. HC, myosin heavy chain; AC, actin; LC, light chains. The myofibril samples have been overloaded so that the light chains, which are present only as ~4% of the total protein, can be clearly distinguished. In scallop myosin the two types of light chains (structural and regulatory) have the same mobility, and desensitization leads to the removal of only the regulatory light chains from the myosin²¹.

Methods. A, QY13 cells containing pLCIIFXMLC or the mutated constructs were grown at 30 °C in 2×TY medium (containing 25 µg ampicillin per ml). At $A_{600\text{ nm}} = 0.8$ the cultures were quickly warmed to 42 °C and maintained at this temperature for 15 min, then grown at 37 °C for 3.5 h. The cell pellets were prepared as described³¹, washed four times with 0.5% Triton X-100, 1 mM EDTA (shown in a)³², dissolved in 6 M urea, 50 mM Tris-HCl, pH 7.5, 0.5 mM MgCl₂, 0.5 mM dithiothreitol (DTT), dialysed against the same solution for 3 h and then applied to a 11×4 cm DEAE-cellulose column (Whatman DE-52) equilibrated in the same solution. The fusion proteins were eluted with a linear gradient (500+500 ml) from 0 to 80 mM NaCl in the 6 M urea solution, and the peak fractions concentrated in an Amicon ultrafiltration cell with a PM-10 membrane (b). The proteins (5 mg ml⁻¹) were dialysed against two changes of 100 mM NaCl, 40 mM Tris-HCl, pH 8.0, 0.5 mM DTT and then digested at 20 °C for 2.5 h with thrombin (enzyme/protein (w/w) ratio 1:250). The digests were diluted with an equal volume of 4 M urea, 25 mM Tris-HCl, pH 7.5, 0.5 mM MgCl₂, 0.5 mM DTT (to reduce the NaCl concentration to 50 mM), loaded onto a 8×4 cm DEAE-cellulose column equilibrated in 2 M urea, 25 mM Tris-HCl, pH 7.5, 0.5 mM MgCl₂, 0.5 mM DTT and the light chains eluted with a linear gradient (400+400 ml) from 50 to 150 mM NaCl in the 2 M urea solution. The light chains were concentrated by ultrafiltration with an Amicon PM-10 membrane (c) and their protein concentrations measured assuming an $E_{280\text{ nm}}^{1\%} = 5.5$ (recovery > 80% of the initial fusion protein). Final yield was 20 mg purified light chain per l of bacterial culture. B, Scallop myofibrils were prepared and desensitized at 27 °C as described previously²¹. The hybrids were prepared by incubating desensitized scallop myofibrils (~2 mg ml⁻¹) in 50 mM NaCl, 4 mM MgCl₂, 5 mM phosphate buffer, pH 7.0, 0.2 mM DTT with the light chains and mutants at a 1.5:1 light chain/myosin heavy chain molar ratio, overnight with stirring at 4 °C. The myofibril-hybrids were washed three times with a ~40-fold excess of the 50 mM NaCl solution to remove the unbound light chains and aliquots taken for gel electrophoresis and for measuring the actin-activated Mg²⁺-dependent ATPase activities shown in Table 1. All the manipulations with the light chain/mutant hybrids were carried out in parallel so that direct comparisons could be made. These assays (shown here and in Table 1) were repeated with three different desensitized scallop myofibril preparations and light chain/mutant light chain binding to these preparations tested under the following conditions: (1) at 0.5 mM, 2 mM or 4 mM Mg²⁺ at 50 mM and 100 mM NaCl; (2) at 50 mM NaCl, 2 mM Mg²⁺ in the presence and absence of 1 mM MgATP and either 0.1 mM Ca²⁺ or 0.5 mM EGTA. Under all the conditions tested similar results to those shown here and in Table 1 were observed.

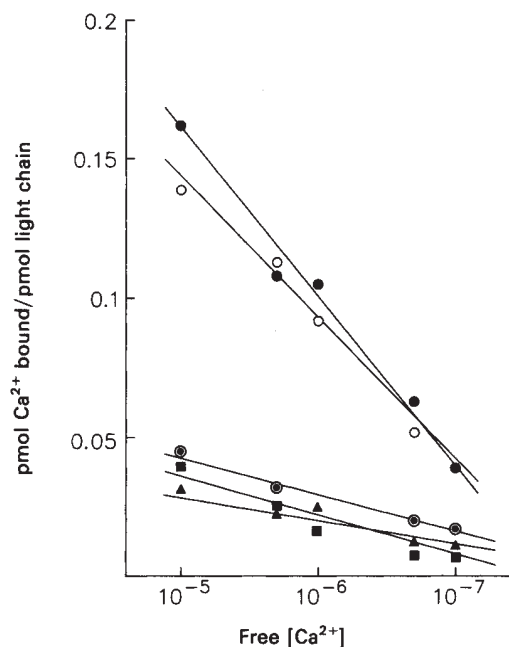
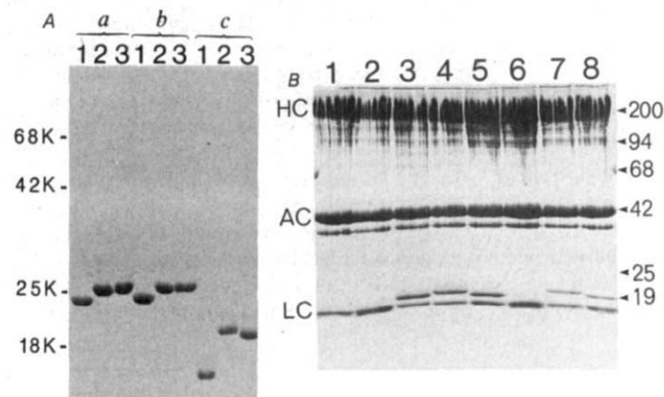


Fig. 3 Ca²⁺-binding ability of the regulatory light chain and mutants. ○, Chicken breast muscle myosin light chain; ●, *E. coli*-synthesized light chain (wild type); ⊙, mutant 463; ▲, mutant 452; ■, mutant 442.

Methods. Ca²⁺ binding was measured using a procedure adapted from that of Kawasaki *et al.*²⁷ using a 96-well Microtiter hybrid dot manifold assembly (BRL, No. 1050 MM) and a nitrocellulose membrane filter (Schleicher & Schuell PH79, 0.1 µm). Aliquots (2–5 µl) containing 4–10 µg of light chains/mutants were applied to the walls of the wells of the manifold and mixed with 250 µl of solution with the appropriate free Ca²⁺ concentration (10⁻⁵–10⁻⁷ M Ca²⁺ in 50 mM NaCl, 0.1 mM MgCl₂, 10 mM imidazole buffer, pH 7.0). These solutions were prepared using an apparent binding constant for Ca/EGTA at 25 °C and pH 7.0 of 3×10⁶ M⁻¹ (ref. 33) and with a final concentration of 5–20 µM ⁴⁵Ca/EGTA (specific activity 200–900 c.p.m. pmol⁻¹). The samples were mixed, incubated at 20 °C for 5 min and rapidly filtered on a vacuum pump. A fresh 250-µl aliquot of the appropriate Ca²⁺ solution was added and incubated at 20 °C for a further 5 min, and then filtered under vacuum. The nitrocellulose filter membrane was dried, wrapped in Saran wrap and autoradiographed at -70 °C for 2 h. The regions of the nitrocellulose membrane containing the light chains/mutants could be identified from the autoradiograph and by the indentation caused by pressure and were excised with a one-hole punch, 'dissolved' in Aquasol-2 (NEN Research Products) and counted. Control wells with no added protein, but washed with the appropriate ⁴⁵Ca/EGTA-containing solutions, were used to correct for background. This procedure was checked by measuring the Ca²⁺ binding of the wild type and mutant light chains by the Hummel-Dryer technique³⁰ at 10⁻⁵ M free Ca²⁺. Ca²⁺ binding measurements were also carried out in the absence of Mg²⁺ and the binding curves obtained were very similar (~5–10% increase in Ca²⁺ binding) to those shown here in the presence of 100 µM Mg²⁺. Mg²⁺ was included to eliminate background low-affinity nonspecific divalent cation binding. Binding measurements were repeated at least three times under each set of conditions.

polypeptide which co-migrated with the chicken skeletal myosin light chain on SDS gels. N-terminal sequence analysis of the cleaved protein gave the sequence Ala-Pro-Lys-Lys-Ala-Lys-Arg, which is identical to the N terminus of the light chain^{10,25} except for the absence of the trimethyl blocking group on the N-terminal alanine residue²⁶. This light chain synthesized in *E. coli* bound to desensitized scallop myofibrils and inhibited the

ATPase activity in an identical manner to the native regulatory light chain of chicken skeletal muscle.

In the divalent cation binding domains of the calcium-binding proteins, the acidic residues at the coordination positions X and -Z are conserved and are believed to bind to Ca²⁺ (refs 3, 4; Fig. 1). To inhibit Ca²⁺ binding in the light chain, we have separately replaced Asp 36 and Asp 47 with Ala at the X and

–Z coordination positions by oligonucleotide-directed mutagenesis. We sequenced a representative clone (463) of the Asp 47 → Ala mutant and showed that it contains no other substitutions. All the 36 clones sequenced in the experiment to produce the Asp 36 → Ala mutant had three additional amino-acid substitutions in the E helix preceding the binding loop (Fig. 1). During the screening for Asp 36 → Ala mutants, we found a clone (442) which contained in addition a deletion of 16 amino acids spanning the phosphorylation site and four further amino-acid substitutions in the N-terminal region (Fig. 1). These mutants were produced in *E. coli*, purified and cleaved with thrombin. As expected, the protein produced from clone 442 was smaller (relative molecular mass (M_r) = 17,000) than the other mutants (M_r = 19,000) (Fig. 2A).

We measured the relative Ca^{2+} binding abilities of the wild-type and mutant light chains at 10^{-7} to 10^{-5} M free Ca^{2+} and found that the Ca^{2+} binding capacities of the mutants were lower than that of the wild-type and native light chains (Fig. 3). All the mutants bound to desensitized scallop myofibrils with an affinity comparable with that of the wild type made in *E. coli* and native chicken myosin regulatory light chains (Fig. 2B). Table 1 shows that when scallop myofibrils were reconstituted with chicken skeletal myosin light chains or those of the wild type made in *E. coli*, the ATPase activity was inhibited both in the presence and absence of Ca^{2+} , whereas the mutant light chains activated the ATPase to a level comparable with that observed with the scallop light chains in the presence of Ca^{2+} . This shows that the mutant light chains have lost the ability to inhibit the interactions of myosin with actin and instead lock the actin-activated myosin ATPase in the 'on' state. Thus, they appear to bind to the myosin heavy chains in the correct positions needed to prevent the decrease in ATPase activity observed with the unregulated myosin heads (see Table 1 for control).

Furthermore, our results show that the first domain in the regulatory light chain really is the site with high affinity for divalent cations. The existence of light chains that differ in their physiological activities, a suitable myosin assay system and the generation of functional mutants in *E. coli* should allow us further to dissect the structural features responsible for myosin regulation.

We thank Drs Max Perutz, Hugh Huxley, Alexander MacLeod, Clive Bagshaw, Jane Mitchell, Sandra Citi, Lesley Clayton and Richard Turner for critically reading the manuscript; Ross Jakes for performing the N-terminal sequence analysis; Dr Peter Esnouf for the gift of factor X and thrombin; and Dr Hans C. Thøgersen for purified thrombin and for pointing out the similarity of the thrombin cleavage site in CIIFXMLC and fibrinogen. K.N. was supported by NIH grant RO1 HL 31461-02, F.C.R. was partially supported by CNPq grant 20.0694, Brazil.

Received 20 February; accepted 8 May 1986.

- Kendrick-Jones, J., Szentkiralyi, E. M. & Szent-Gyorgyi, A. G. *J. molec. Biol.* **104**, 747–775 (1976).
- Baba, M. L., Goodman, M., Berger-Cohn, J., DeMaille, J. G. & Matsuda, G. *Molec. biol. Evol.* **1**, 442–455 (1984).
- Kretsinger, R. H. *CRC Crit. Rev. Biochem.* **8**, 119–174 (1980).
- Kretsinger, R. H. & Nockolds, C. E. *J. biol. Chem.* **248**, 3313–3326 (1973).
- Babu, Y. S. *et al. Nature* **315**, 37–40 (1985).
- Herzberg, O. & James, M. N. G. *Nature* **313**, 653–659 (1985).
- Sundaralingam, M. *et al. Science* **227**, 945–948 (1985).
- Collins, J. H. *Nature* **259**, 699–700 (1976).
- Kendrick-Jones, J. & Jakes, R. in *Boehringer International Symposium on Myocardial Failure* (eds Rieker, G., Weber, A. & Goodwing, J.) 28–40 (Springer, Berlin, 1976).
- Matsuda, G., Suzuyama, Y., Maita, T. & Umegane, T. *FEBS Lett.* **84**, 53–56 (1977).
- Bagshaw, C. R. & Kendrick-Jones, J. *J. molec. Biol.* **130**, 317–336 (1979).
- Bagshaw, C. R. *Biochemistry* **16**, 59–67 (1977).
- Craig, R. *et al. J. molec. Biol.* **140**, 35–55 (1980).
- Vibert, P. & Craig, R. *J. molec. Biol.* **157**, 299–319 (1982).
- Flicker, P. F., Wallimann, T. & Vibert, P. *J. molec. Biol.* **169**, 723–741 (1983).
- Adelstein, R. S. & Eisenberg, E. A. *Rev. Biochem.* **49**, 921–956 (1980).
- Kendrick-Jones, J. *et al. J. molec. Biol.* **165**, 139–162 (1983).
- Leavis, P. C. & Gergely, J. *CRC Crit. Rev. Biochem.* **16**, 235–305 (1984).
- Persechini, A., Stull, J. T. & Cooke, R. *J. biol. Chem.* **260**, 7951–7954 (1985).
- Szent-Gyorgyi, A. G., Szentkiralyi, E. M. & Kendrick-Jones, J. *J. molec. Biol.* **74**, 179–203 (1973).
- Chantler, P. D. & Szent-Gyorgyi, A. G. *J. molec. Biol.* **138**, 473–492 (1980).

- Kendrick-Jones, J. *et al. in Basic Biology of Muscles: A Comparative Approach* (eds Twarog, B. M., Levine, R. J. C. & Dewey, M. M.) 255–272 (Raven, New York, 1982).
- Bagshaw, C. R. & Kendrick-Jones, J. *J. molec. Biol.* **140**, 411–433 (1980).
- Blombäck, M., Blombäck, B., Mammen, E. F. & Pradsad, A. S. *Nature* **218**, 314–318 (1968).
- Reinach, F. C. & Fischman, D. A. *J. molec. Biol.* **181**, 411–422 (1985).
- Hendry, G. D. *et al. FEBS Lett.* **144**, 11–15 (1982).
- Kawasaki, H., Kasai, H. & Okuyama, T. *Analyt. Biochem.* **148**, 297–302 (1985).
- Nagai, K. & Thøgersen, H. C. *Nature* **309**, 810–812 (1984).
- Carter, P., Bedouelle, H. & Winter, G. *Nucleic Acids Res.* **13**, 4431–4443 (1985).
- Murray, N., Bruce, S. A. & Murray, K. *J. molec. Biol.* **132**, 493–505 (1979).
- Nagai, K., Perutz, M. F. & Poyart, C. *Proc. natn. Acad. Sci. U.S.A.* **82**, 7252–7255 (1985).
- Marston, F. A. O. *et al. Biotechnology* **2**, 800–804 (1984).
- Marston, S. B. *Prog. Biophys. molec. Biol.* **41**, 1–41 (1983).

Enzymatic activity of the conserved core of a group I self-splicing intron

Jack W. Szostak

Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

Splicing of the *Tetrahymena* ribosomal intron was first studied by Cech *et al.*^{1,2}, who subsequently demonstrated that the intron RNA catalyses its own excision from a primary transcript to yield mature ribosomal RNA^{3–5}. This intron shares several short conserved sequences and a common secondary structure with several other introns^{6–9}, some of which have also been shown to self-splice^{10,11}. Here I show that the conserved core of the *Tetrahymena* intron can act in *trans* to catalyse the sequence-specific cleavage and addition of guanosine to a separate RNA substrate.

Cech *et al.* proposed that self-splicing occurs by a series of *trans*-esterification reactions in which the total number of phosphodiester bonds is conserved^{4,12,13}. Each step involved the concerted breakage and formation of a specific phosphodiester bond, catalysed by a single phosphoester transferase activity contained within the intron RNA itself. In the first step, the 3' hydroxyl of a guanosine^{12,14} attacks the phosphodiester bond between the 5' exon and the intron. That bond is broken and a new one is formed between the guanosine and the 5' end of the intron, thus conserving the total number of phosphodiester bonds. The second step is another exchange reaction; in this case the phosphodiester bond between the 3' end of the intron and the 3' exon is broken and a new one is formed between the 5' and 3' exons. In the third step, the intron circularizes itself in yet another *trans*-esterification reaction.

Examination of potential secondary structures shows that all three of these reactions may take place in a very similar structural context (Fig. 1). In each case, the reaction is between one strand of a short region of double-stranded RNA and a guanosine residue which can be either free G or the 3' nucleotide of a chain. Steps one and three are completely analogous, while step two is the reverse reaction. The similarity of the substrates for the three reactions strengthens the view that the entire splicing process is catalysed by a single enzyme activity. If the substrate was defined by local sequence and structure, and not by the overall folding of the intron, it seemed likely that the enzyme might also be capable of catalysing the reaction on an exogenously supplied template. Figure 2 shows the structure of the conserved core of the *Tetrahymena* rRNA intron, drawn according to Davies *et al.*⁶. This segment of the intron contains the conserved P, Q, R and S sequences of Davies *et al.*, but not the flanking secondary structures. The secondary structure of the core is supported by a considerable degree of phylogenetic comparison^{6,9}, and some structural analysis^{15,16}.

The hypothesis that this conserved core contains the enzymatic activity was tested by constructing a series of plasmids (Fig. 3) that could be transcribed *in vitro* by T7 RNA polymerase to yield RNA molecules that could be tested for self-splicing or *trans*-esterification activity. The first experiment involved the insertion, downstream of a T7 RNA polymerase promoter, of

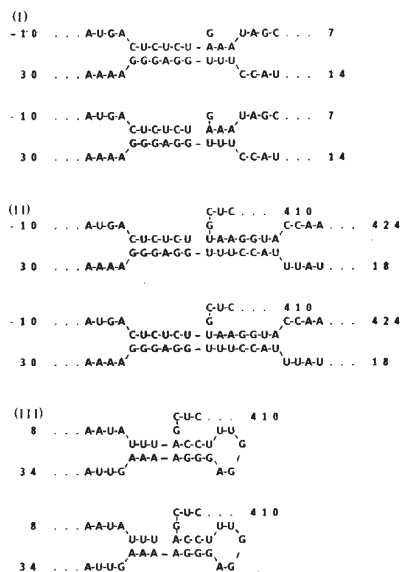


Fig. 1 Substrates for the RNA enzyme. The first *trans*-esterification reaction (I) is between a free G residue and an A residue that is part of a stem-loop structure. The double-stranded stem is formed between the exon-intron junction sequence, and the internal guide sequence^{6,9}. Release of the G-intron allows a more extensive pairing between the internal guide sequence and the 3' intron-exon sequence (II). The intron ends in G, which is placed in the same position as the incoming G in reaction I. Catalysis of the reverse reaction then results in loss of the intron and joining of the 5' and 3' exons. Circularization of the intron (III) takes place in a similar sequence and secondary structure context. The first nucleotide of the intron is +1 in the sequence coordinate system.

a DNA restriction fragment spanning the entire intron. As expected, RNA transcribed *in vitro* from this DNA and gel-purified was able to perform all three steps of the self-splicing process when incubated in buffer containing Mg^{2+} and guanosine. A series of RNAs truncated at different 3' positions was generated by cleavage of the DNA template with restriction endonucleases, followed by T7 RNA polymerase runoff transcription. The enzyme *ScaI* cuts just inside the 3' end of the intron, and templates cleaved with this enzyme yield RNA molecules lacking the 3' intron-exon junction. Remarkably, these transcripts still splice, apparently finding new acceptor splice sites within the intron sequence. The enzyme *NheI* cuts the DNA 20 nucleotides 3' of the conserved core, while *BglII* cuts within the core region. Transcription of *NheI*-cut DNA yielded enzymatically active RNA, as incubation of the gel-purified RNA with guanosine in splicing buffer resulted in the appearance of smaller RNA products. The RNA resulting from transcription of *BglII*-cut DNA was inert. All subsequent experiments were therefore done with transcripts of *NheI*-cut DNA.

The extent of 5' sequences required for enzymatic activity was determined by constructing internal deletions. Since the RNAs synthesized from these templates lack all of the sequences at which splicing normally takes place, an exogenous substrate was added to assay for *trans*-esterification activity. The three plasmids used to generate the shortened catalytic RNAs are shown in Fig. 3; pLG8 and pLG9 delete intron sequences to within 84 and 51 nucleotides of the conserved core (the first nucleotide of the P3 segment, GACCGUCA, in Fig. 2), while in pSZ243, all intron sequences farther than 9 nucleotides from the core are deleted. These catalytic RNAs were tested on a substrate prepared by T7 RNA polymerase transcription of pSZ241 cut with *DraI*. This generates a 119-nucleotide RNA containing the sequences that are the substrate for the first *trans*-esterification reaction (Fig. 1). The substrate RNA begins

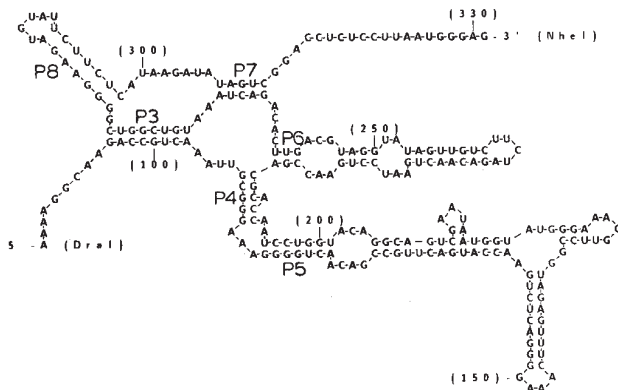


Fig. 2 Secondary structure of the conserved core of the RNA enzyme. All of the group I introns are characterized by conserved sequence elements and a conserved secondary structure. The conserved pairing regions P3, P4, P5, P6, P7 and P8 are labelled in accordance with the model of Davies *et al.*⁶. The nucleotide coordinates are numbered from the beginning of the intron. The segment of the *Tetrahymena* rRNA intron illustrated here corresponds to the T7 transcript of plasmid pSZ243 digested with *NheI*. Thus, reaction (I) (Fig. 1) involves the action of this catalytic core on sequences 100 nucleotides upstream, and reaction (II) involves substrate sequences both upstream and downstream from the conserved core.



Fig. 3 Plasmids and transcripts used in the present study. The location of the intron relative to certain restriction sites is indicated by the heavy-lined segment. Plasmid pSZ241 contains the entire intron flanked by exon sequences. It was constructed by inserting the *HinI* fragment of pTre1 (obtained from E. H. Blackburn) into the *SmaI* site of the T7 transcription vector pST53 (obtained from S. Tabor). Enzymatically active RNAs were generated by transcription of *BamHI*, *ScaI* or *NheI* digests of this DNA with T7 RNA polymerase. Plasmid pLG8 was generated by partial digestion of pSZ241 with *SspI*, complete digestion with *SmaI*, followed by blunt-end ligation. Plasmid pLG9 was generated by digestion of pSZ241 with *SmaI* and *SphI*, treatment with T4 polymerase to generate blunt ends, followed by blunt-end ligation. pSZ243 was constructed by blunt-end ligation of the gel-purified *DraI* fragment of pSZ241 into *SmaI*-cut pSZ241, followed by *NheI* digestion and recircularization. Substrate RNA was made by T7 transcription of *DraI*-cut pSZ241 DNA. All transcripts were purified by acrylamide-urea gel electrophoresis.

with 25 nucleotides of vector sequences, followed by 32 nucleotides of the 5' exon and the first 62 nucleotides of the intron. Activity was assayed by mixing catalytic RNA with substrate RNA in the presence of labelled GTP. The activity of the full-length intron RNA and of the three deletion derivatives was assayed (Fig. 4). All of the catalytic RNAs were able to cleave the substrate RNA at the exon-intron junction. The major labelled product is formed by the addition of GTP to the 5' end of the intron sequence. No incorporation of labelled G is observed in the absence of RNA enzyme. Since the pSZ243-*NheI* transcript contains only the conserved core sequence flanked by 10-20 nucleotides at each end, the conserved core does in fact contain the enzyme activity responsible for the *trans*-esterification reactions.

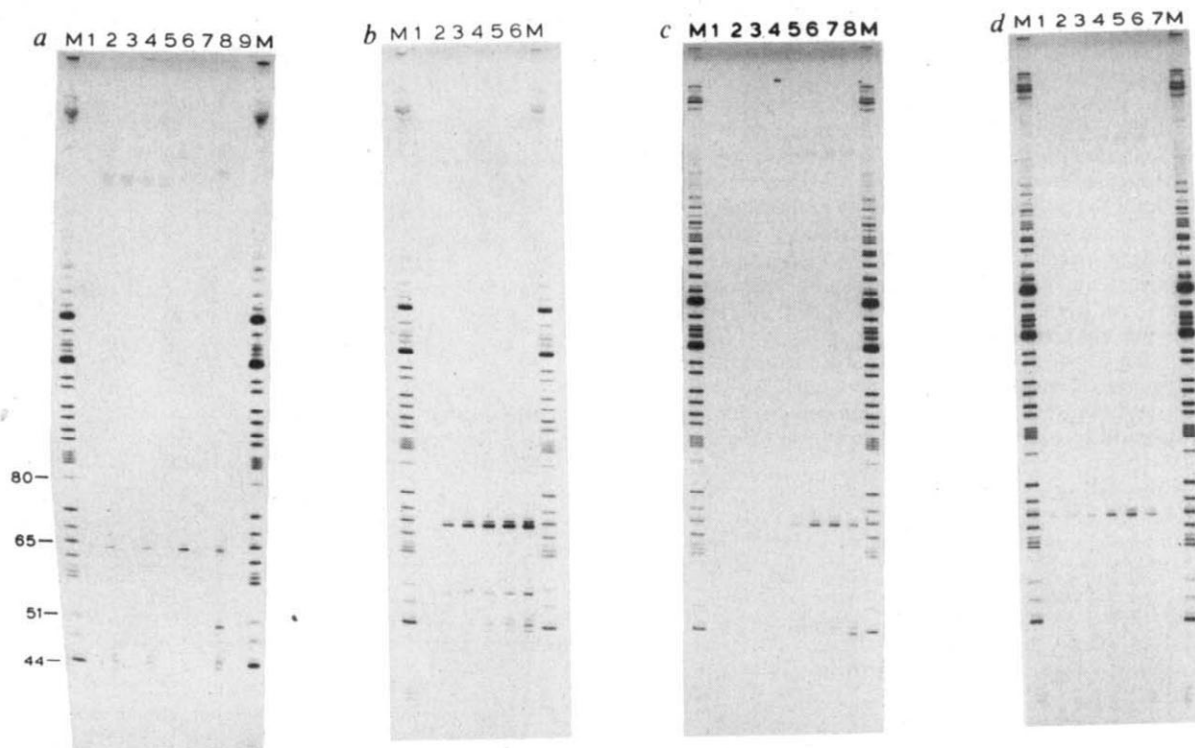


Fig. 4 Products of RNA-catalysed *trans*-esterification. **a**, Activity of deletion derivatives. Lane 1 contains substrate (transcript of *Dra*I-cut pSZ241) but no enzyme (catalytic RNA). Lanes 2, 4, 6 and 8 contain enzyme and substrate; lanes 3, 5, 7 and 9 contain enzyme but no substrate. Lanes 2 and 3 show the addition of G and cleavage of substrate by the T7 transcript of *Nhe*I-cut pSZ241. Lanes 4 and 5 contain transcript of *Nhe*I-cut pLG8, lanes 5 and 6 contain the pLG9-*Nhe*I transcript, and lanes 7 and 8 contain the pSZ243-*Nhe*I transcript. The major product runs just below the 65-nucleotide marker, and corresponds to the GTP-intron sequence addition product. All reactions were done in the presence of 10 mM NH_4Cl , 20 mM MgCl_2 , 30 mM Tris-HCl pH 7.4 and 0.2 mM aurin tricarboxylic acid. Incubations were for 30 min at 58 °C. Catalytic RNAs were assayed at a concentration of ~50 nM, substrate RNAs were present at 0.5 μM , and α -labelled GTP at 2.5 μM . Size markers (Lanes M) are the 'A'-lane of a dideoxy-terminated sequence analysis reaction using a primer corresponding to the first 17 nucleotides of the substrate transcript, and DNA of pSZ241 as template. Sizes are 31, 32, 33, 44, 45, 48, 51, 58, 59, 60, 62, 65, 66, 68, 72 and 80 nucleotides for the region of interest. **b**, Time course of reaction. Aliquots were removed at 0, 15, 30, 45, 60 and 90 min (lanes 1-6) from a reaction of the enzymatically active pSZ243-*Nhe*I transcript, with α -labelled GTP and the pSZ241-*Dra*I transcript as substrates. Reaction conditions were as for **a**. **c**, Optimal Mg^{2+} concentration. Reactions were as in **b** for 30 min, except that the Mg^{2+} concentration was varied. Reactions in lanes 1-8 contained 0, 2, 5, 10, 20, 30, 50 and 100 mM Mg^{2+} , respectively. **d**, Temperature optimum. Reactions were as in **b** for 15 min, except that the temperature was varied. Lanes 1-7 show reactions incubated at 42, 46, 50, 54, 58, 62 and 66 °C, respectively.

The initial experiments used reaction conditions similar to those of Bass and Cech¹⁴. Characterization of the core enzyme, however, revealed some changes in optimum reaction conditions. The optimal Mg^{2+} concentration had increased from the 5 mM reported for the intact intron, to 20-40 mM (Fig. 4). This may reflect the difference between the intramolecular reaction of the intact intron and the intermolecular reaction, which requires a close interaction between separate substrate and enzyme molecules. This is supported by the effect of the volume excluder polyethylene glycol, which had a strong rate-enhancing effect at low Mg^{2+} concentration, but very little effect at a higher Mg^{2+} concentration. The temperature optimum of the reaction was a remarkably high 58 °C (Fig. 4), which was particularly surprising in view of the relatively short double-stranded region in the substrate stem loop and the short base-paired regions characteristic of the core enzyme itself. It is likely that the structure of the enzyme is stabilized by numerous tertiary interactions, and that the enzyme in turn stabilizes the structure of the substrate.

The major products of the *trans*-esterification reaction clearly correspond to those expected by sequence-specific cleavage and GTP addition at the exon-intron boundary. Several minor products are visible, and may reflect a decreased sequence specificity of the small enzyme RNA. The identity of the major product

of the reaction was confirmed by examining the products of cleavage of 5'-end labelled RNA and uniformly labelled RNA (data not shown). The 3' fragment is not seen when 5' labelled RNA and cold G are used. With uniformly labelled RNA, the mobility of the 3' fragment is approximately 1 nucleotide faster if GTP rather than G is used as co-substrate, as expected from the extra negative charge contributed by the triphosphate moiety. The expected 57-nucleotide 5' fragment is seen with both 5' and uniformly labelled substrate RNA, but in submolar yield relative to the 3' fragment, probably because of splicing onto other substrate or enzyme sequences. In addition, other fragments are seen when labelled substrate RNA is used, possibly due to an inherent hydrolytic activity of the enzyme.

The observation that the self-splicing intron can be dissociated into an enzymatically active core and distinct substrate sequences should facilitate the genetic and biochemical analysis of both enzyme and substrate. Experiments with modified substrates should define the aspects of the substrate that are recognized by the enzyme. It should be a straightforward matter to define the minimal sequences required for enzymatic activity. Mutations in the enzyme which have specific effects on the binding of the RNA substrate or the guanosine co-substrate, or on the rate or nature of the reaction, should be informative with respect to the mechanism of catalysis. In addition, the constraints

that such information will place on enzyme structure and enzyme-substrate interactions will facilitate the development of models of the tertiary structure of the enzyme.

It is widely thought that RNA catalysis played an important part early in the evolution of life, before the evolution of protein synthesis. Sharp has suggested that early RNA replicases worked by splicing together short oligonucleotides in a template-directed manner¹⁷. An RNA polymerase-like activity of the *Tetrahymena* intron has recently been reported¹⁸. Mutations that lead to a loss in sequence specificity might result in a more general polymerase activity. It may therefore be possible to design an RNA replicase made of RNA by introducing a relatively small number of changes into the *Tetrahymena* intron core.

I thank Stan Tabor and Charles Richardson for gifts of T7 RNA polymerase, Laura Glew and Anne Gerber for technical assistance, and Andrew Murray for comments on the manuscript. This work was supported by a grant from Hoechst AG.

Received 25 February; accepted 15 April 1986.

1. Zaug, A. J. & Cech, T. R. *Cell* **19**, 331-338 (1980).
2. Grabowski, P. J., Zaug, A. J. & Cech, T. R. *Cell* **23**, 467-476 (1981).
3. Kruger, K. et al. *Cell* **31**, 147-157 (1982).
4. Zaug, A. J., Grabowski, P. J. & Cech, T. R. *Nature* **301**, 578-583 (1983).
5. Zaug, A. J., Kent, J. R. & Cech, T. R. *Science* **224**, 574-578 (1984).
6. Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A. & Scazzocchio, C. *Nature* **300**, 719-724 (1982).
7. Michel, F. & Dujon, B. *EMBO J.* **2**, 33-38 (1983).
8. Cech, T. R. et al. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3903-3907 (1983).
9. Waring, R. B., Scazzocchio, C., Brown, T. A. & Davies, R. W. *J. molec. Biol.* **167**, 595-605 (1983).
10. Garriga, G. & Lambowitz, A. M. *Cell* **38**, 631-641 (1984).
11. Horst, G. v. d. & Tabak, H. F. *Cell* **40**, 759-766 (1985).
12. Cech, T. R., Zaug, A. R. & Grabowski, P. J. *Cell* **27**, 487-496 (1981).
13. Cech, T. R. *Cell* **34**, 713-716 (1983).
14. Bass, B. & Cech, T. R. *Nature* **308**, 820-826 (1984).
15. Inoue, T. & Cech, T. R. *Proc. natn. Acad. Sci. U.S.A.* **82**, 648-652 (1985).
16. Tanner, N. K. & Cech, T. R. *Nucleic Acids Res.* **13**, 7759-7779 (1985).
17. Sharp, P. A. *Cell* **42**, 397-400 (1985).
18. Zaug, A. J. & Cech, T. R. *Science* **231**, 470-475 (1986).

Mechanism of recognition of the 5' splice site in self-splicing group I introns

Gian Garriga*, Alan M. Lambowitz*, Tan Inoue† & Thomas R. Cech‡

* Edward A. Doisy Department of Biochemistry, St Louis University School of Medicine, St Louis, Missouri 63104, USA

† Salk Institute, La Jolla, California 92037, USA

‡ Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309, USA

Group I introns include many mitochondrial ribosomal RNA and messenger RNA introns and the nuclear rRNA introns of *Tetrahymena* and *Physarum*¹⁻⁶. The splicing of precursor RNAs containing these introns is a two-step reaction. Cleavage at the 5' splice site precedes cleavage at the 3' splice site, the latter cleavage being coupled with exon ligation⁷⁻⁹. Following the first cleavage, the 5' exon must somehow be held in place for ligation. We have now tested the reactivity of two self-splicing group I RNAs, the *Tetrahymena* pre-rRNA and the intron 1 portion of the *Neurospora* mitochondrial cytochrome *b* (*cob*) pre-mRNA, in the intermolecular exon ligation reaction (splicing in *trans*) described by Inoue et al.¹⁰. The different sequence specificity of the reactions supports the idea that the nucleotides immediately upstream from the 5' splice site are base-paired to an internal, 5' exon-binding site, in agreement with RNA structure models proposed by Davies and co-workers^{2,3} and others^{4,5,11}. The internal binding site is proposed to be involved in the formation of a structure that specifies the 5' splice site and, following the first step of splicing, to hold the 5' exon in place for exon ligation.

When the dinucleotide CpU or pCpU is incubated with the *Tetrahymena* pre-rRNA, the RNA is cleaved exactly at its 3'

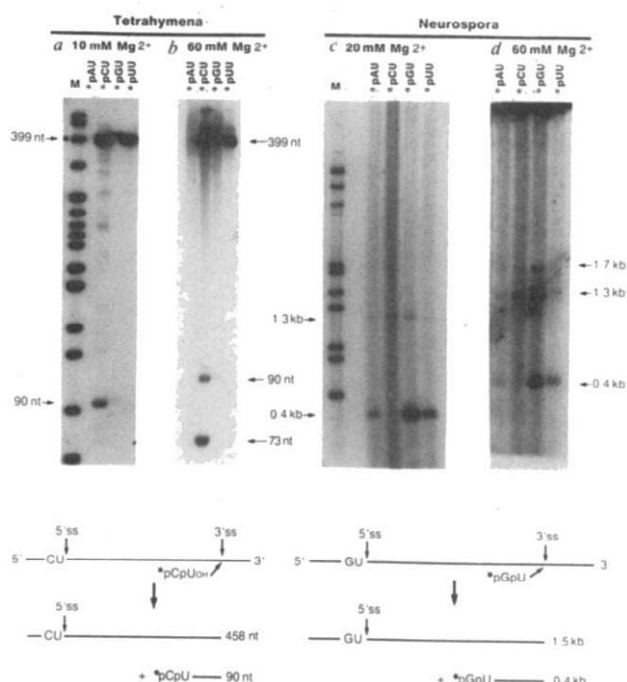


Fig. 1 Intermolecular exon ligation. *Tetrahymena* pre-rRNA was incubated with ³²P-pApU, ³²P-pCpU, ³²P-pGpU or ³²P-pUpU in *Tetrahymena* reaction conditions (10 mM Mg²⁺) (a) or *Neurospora* reaction conditions (60 mM Mg²⁺) (b). *Neurospora cob* pre-mRNA was similarly incubated in *Neurospora* reaction conditions (20 mM Mg²⁺) (c) or 60 mM Mg²⁺ (d). *Tetrahymena* and *Neurospora* reaction conditions are defined below. The diagrams at the bottom show *trans*-splicing reactions for *Tetrahymena* and *Neurospora* RNAs (not drawn to scale).

Methods. *Tetrahymena* pre-rRNA was prepared by transcription of pT7TT1A3, a phage T7 promoter version of the SP6 promoter-containing plasmid pSPTT1A3 (ref. 20). *Neurospora cob* intron 1 RNA was prepared by transcription of pSP64-H2a, using SP6 RNA polymerase at 40 °C in the absence of spermidine¹². Dinucleotides ApU, CpU, GpU and UpU (Pharmacia, P-L Biochemicals) were 5' end-labelled using polynucleotide kinase and [γ -³²P]ATP¹⁰. Each labelled dinucleotide ran as a single band on a 20% polyacrylamide gel; the dinucleotides were not significantly cross-contaminated by this criterion. Splicing reactions were carried out by incubating 1.17 pmol of the *Neurospora* transcript or 0.03 pmol of the *Tetrahymena* transcript with ³²P-pCpU, ³²P-pApU, ³²P-pGpU or ³²P-pUpU (12 nmol for *Neurospora*, 0.2 nmol for *Tetrahymena*) for 1 h in 5 μ l of reaction media. The experiment in a was carried out under *Tetrahymena* reaction conditions: the reaction media contained 200 mM NH₄Cl, 10 mM MgCl₂, 50 mM EPPS (*N*-2-hydroxyethylpiperazine-*N'*-2'-propanesulphonic acid)-HCl pH 7.5, and the splicing reactions were carried out at 30 °C. These conditions differ somewhat from those used previously¹⁰. The experiments in b, c and d were carried out under *Neurospora* reaction conditions: the reaction media contained 50 mM NH₄Cl, 20 or 60 mM MgCl₂, and 50 mM EPPS-HCl pH 7.5, and the splicing reactions were carried out at 37 °C. The reactions were terminated by addition of EDTA (100 mM), and RNAs were precipitated with ethanol three times before analysis. *Tetrahymena* RNAs (a and b) were analysed on an 8% polyacrylamide (20:1, acrylamide:bisacrylamide) gel containing 7 M urea, 90 mM Tris-borate, 2.5 mM EDTA. *Neurospora* RNAs (c and d) were glyoxalated and analysed on a 1.6% agarose gel²¹; ethidium bromide staining confirmed that equal amounts of RNA had been loaded on each lane. The gels were dried and analysed by autoradiography.

splice site and the dinucleotide is covalently ligated to the 3' exon¹⁰. This reaction can be explained as an intermolecular version of exon ligation, because it occurs with oligonucleotides whose sequence resembles that of the 5' exon (CUCUCU), and their site of attachment is identical to that of exon ligation. Except for the terminal uridine residue, the sequence preceding

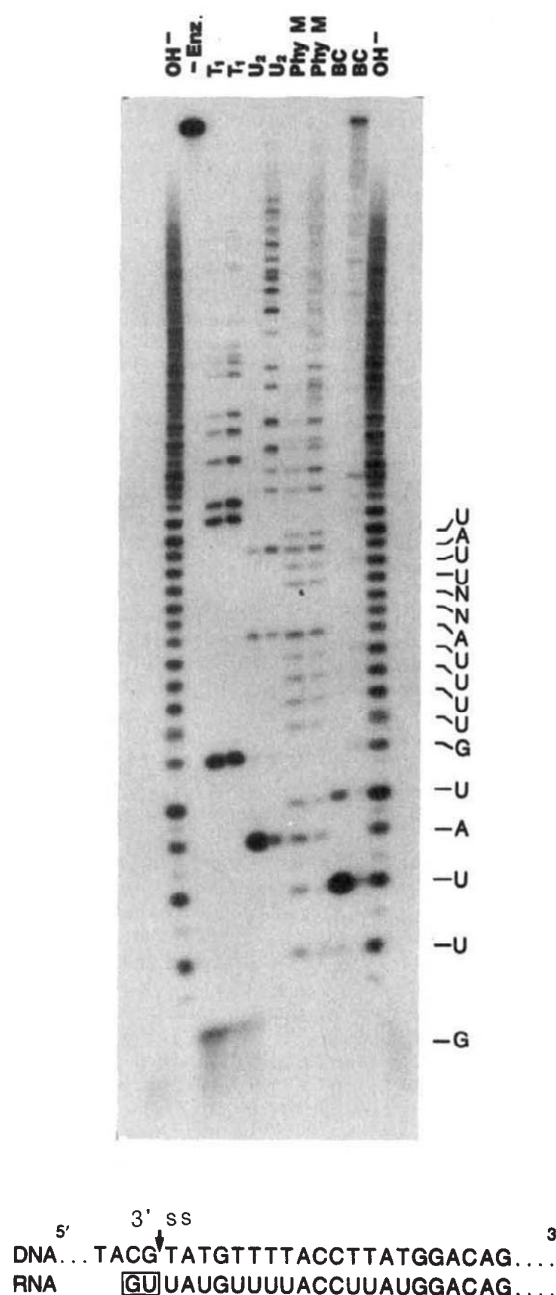


Fig. 2 RNA sequencing of the ^{32}P -pGpU-labelled 0.4-kb RNA. The 0.4-kb *Neurospora* RNA (conditions of Fig. 1d, 3 h incubation) was purified by gel electrophoresis. The RNA was then sequenced by partial hydrolysis with ribonucleases and alkali (OH^-)²²⁻²⁴. Ribonuclease T_1 is specific for G residues, ribonuclease U_2 for A residues, ribonuclease Phy M for A and U residues, and *Bacillus cereus* (BC) ribonuclease for C and U residues. '-Enz.' shows RNA incubated under the conditions used for RNase T_1 , but in the absence of enzyme. 3' ss, 3' splice site; the boxed GU represents RNA sequencing reactions were analysed on a gel containing 22% polyacrylamide (20:1, acrylamide:bisacrylamide), 7 M urea, 90 mM Tris-borate, 2.5 mM EDTA.

the 5' splice site is not conserved among group I introns. For example, among other group I introns that have been found to be self-splicing *in vitro*, the first intron of the *Neurospora cob* gene is preceded by CUGGU (ref. 12) the yeast rRNA intron by AGGGAU (ref. 13) and intron 5a of the yeast cytochrome oxidase subunit 1 gene by AUGAUU (refs 14, 15). If each of

these sequences defines a 5' splice site because it pairs with a sequence within the intron, the intermolecular exon ligation reaction should show a preference for oligonucleotides of different sequence. On the other hand, if the reaction of the *Tetrahymena* pre-rRNA with CpU reflects its interaction with one of the sequence elements that is conserved among group I introns, or some unanticipated higher nucleophilicity of CpU compared with the other dinucleotides, then CpU might be the most reactive dinucleotide towards all the precursor RNAs.

Truncated versions of the *Tetrahymena* pre-rRNA and *Neurospora cob* pre-mRNA were synthesized *in vitro* using either SP6 or T7 RNA polymerase. The transcripts were incubated with ^{32}P -labelled dinucleotides (^{32}P -pNpU_{OH}) under either *Tetrahymena* or *Neurospora* conditions (see Fig. 1 legend) in reaction media containing different concentrations of MgCl_2 . Figure 1a and b shows incubation of the *Tetrahymena* RNA with the four labelled dinucleotides under optimal splicing conditions for that intron (*Tetrahymena* conditions, 10 mM MgCl_2) or under *Neurospora* conditions in a reaction medium containing 60 mM MgCl_2 . In both cases, incubation with ^{32}P -pCpU resulted in labelling of a 90-nucleotide RNA band that is the product of intermolecular ligation of pCpU to the 3' exon. (Size was measured as 92 nucleotides relative to DNA markers; the size of 90 nucleotides is based on sequence of the DNA.) The identity of this band was confirmed by partial digestion of the 90-nucleotide (90-nt) band with RNase T_1 (data not shown). A number of other bands were also labelled, particularly under the high- Mg^{2+} conditions. The 399-nt band labelled by pCpU and pUpU was shown previously to result from reopening of the circular intervening sequence (IVS)¹⁶. The band slightly larger than the IVS labelled by pGpU and the 73-nt band labelled by pCpU under high- Mg^{2+} conditions were not observed previously with other *in vitro* transcripts incubated under somewhat different conditions¹⁰; their identities are unknown. Under both sets of conditions, the *trans*-splicing reaction showed a strong preference for pCpU. We observed no *trans*-splicing with pApU or pUpU, and incubation with pGpU resulted only in very inefficient labelling of a band of ~90 nt (<5% of the reaction with pCpU; Fig. 1a).

Trans-splicing of the *Neurospora* RNA showed the different specificity predicted for the interaction between the dinucleotide and the putative 5' exon-binding site. Figure 1c and d shows incubations of the *Neurospora* RNA with the four labelled dinucleotides in solutions containing 20 mM MgCl_2 and 60 mM MgCl_2 , respectively. In both reaction media, incubation with pGpU gave a prominent labelled RNA of the size expected for the 3' exon (0.4 kilobases (kb); Fig. 1c,d). The identification of this RNA as the *trans*-splicing product is confirmed below. We observed no *trans*-splicing with pCpU, which is the most reactive dinucleotide for the *Tetrahymena* RNA. However, incubation with pApU and pUpU gave about 10% and 20%, respectively, of the extent of reaction with pGpU. The relative reactivities of the four dinucleotides in the putative *trans*-splicing reaction were roughly the same in reaction media containing 10, 20, 30 and 60 mM Mg^{2+} . Considered together, the results show that the *Neurospora trans*-splicing reaction has the specificity predicted for base pairing with the 5' exon-binding site, but that the specificity appears to be somewhat less stringent than that for the *Tetrahymena* RNA. The *trans*-splicing reaction for *Neurospora* RNA also seems to be less efficient than that for the *Tetrahymena* RNA and requires higher concentrations of precursor RNA and dinucleotide (see Fig. 1 legend). The inefficiency of the specific reaction with pGpU necessitated loading more RNA per lane, which may make competing reactions with pApU and pUpU appear more prominent. We note, however, that pCpU shows no *trans*-splicing reaction whatsoever, indicating that this dinucleotide must be somehow incompatible with the exon-binding site.

To confirm that intermolecular exon ligation had in fact occurred for the *Neurospora* RNA, the 0.4-kb RNA labelled by

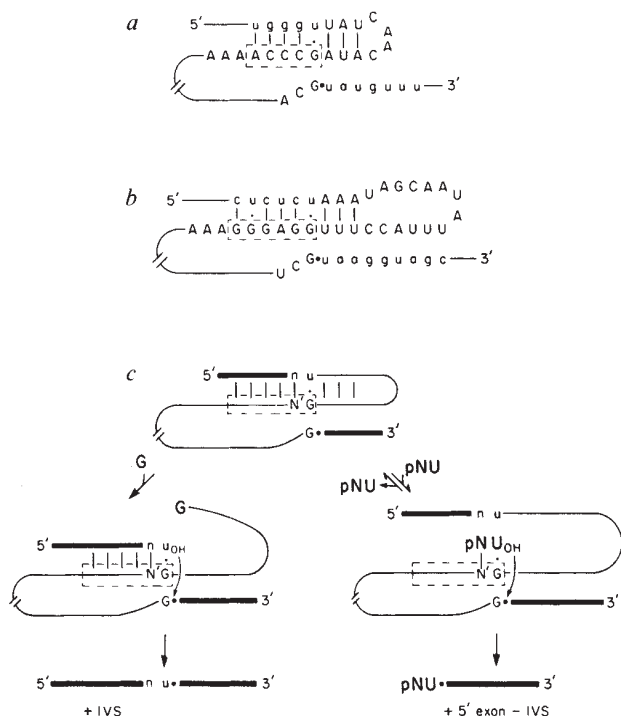


Fig. 3 Internal binding site model accounts for the sequence specificity of the exon ligation reaction. Structures of *N. crassa* *cob* pre-mRNA¹² (a) and *T. thermophila* pre-rRNA^{3,5} (b) are shown. Lower case letters, exons; capital letters, introns; ●, 3' splice site. Boxed nucleotides indicate the portion of the internal guide sequence proposed to pair with the 5' exon^{2,3}. The pairing of the 3' exon with an adjacent portion of the internal guide sequence, which is not supported by mutation analysis²⁵, has been omitted from the diagrams. The interaction that holds the 3' splice site in place is therefore left unspecified. c, Generalized model for the role of the internal binding site. Left, normal exon ligation; right, intermolecular ligation of dinucleotides. The dinucleotide competes with the 5' exon of the intact precursor RNA for binding to the internal guide sequence. When the dinucleotide binds, its free 3'-hydroxyl group is directed towards the phosphorus atom of the phosphate at the 3' splice site in such a way as to facilitate nucleophilic attack. The result is covalent attachment of the dinucleotide to the 3' exon, an intermolecular version of exon ligation.

³²P-pGpU was purified and its sequence determined as shown in Fig. 2. The sequence confirmed that ³²P-pGpU had become covalently attached to the 3' exon by a phosphodiester linkage. The sequence of the 3' exon was the same as determined by Helmer-Citterich *et al.*¹⁷ and Burke *et al.*¹⁸.

Incubation of the *Neurospora* RNA with all four dinucleotides resulted in labelling of two larger bands of 1.7 and 1.3 kb, in addition to the 0.4-kb band. These larger bands are more prominently labelled under the high-Mg²⁺ conditions (Fig. 1c,d). Their sizes are those expected for the excised intron (1.3 kb) and intron plus 3' exon (1.7 kb). Because *Neurospora cob* intron 1 was not observed to undergo cyclization under any conditions tested previously¹², labelling of an intron-sized RNA by circle reopening¹⁶ was not anticipated. One possibility, therefore, is that labelling of the 1.3- and 1.7-kb bands reflects inefficient self-splicing initiated by the dinucleotide substituting for the normal guanosine cofactor. The *Tetrahymena* RNA has been reported to show inefficient splicing with nucleoside triphosphates other than GTP⁹. Again, these competing reactions may appear more prominent here because of the inefficiency of the *trans*-splicing reaction with the *Neurospora* intron.

In the previous study of *trans*-splicing with the *Tetrahymena* intron, the two trinucleotides tested, UCU and UUU, had the

relative reactivities predicted for base-pairing with the putative 5' exon-binding site¹⁰. In analogous experiments, the trinucleotides GGU, AGU and GAU were tested with the *Neurospora* RNA, and in each case the extent of reaction was greater than with the dinucleotide GU (data not shown). The increased reactivity of GGU might be explained on the basis of increased base-pairing interactions with the active site, while the reactivity of the other two trinucleotides might be explained as a combination of base-pairing and increased base-stacking interactions. Any such interpretation was clouded, however, by the unexpected finding that in 10 mM MgCl₂ the *Tetrahymena* RNA reacted efficiently with GGU and GAU, and to a smaller extent with AGU, to give RNAs of the size expected for *trans*-splicing to the 3' exon (91 nt). Partial RNA sequence analysis with RNase T₁ confirmed that GGU was accurately spliced to the 3' exon (data not shown). Thus, although Watson-Crick base-pairing with the 5' exon-binding site may be an important determinant of the *trans*-splicing reaction, additional interactions must also contribute to determining the extent of reaction with trinucleotides. Base stacking or non-Watson-Crick hydrogen-bonding might allow certain trinucleotides to interact with the normal binding site. Alternatively, the trinucleotide might interact with other, at present unidentified, binding sites within the intron.

Although the findings for trinucleotides were unexpected, our results show that *Neurospora cob* intron 1 RNA, like the *Tetrahymena* RNA, undergoes a specific *trans*-splicing reaction with dinucleotides. If the ligation of dinucleotides to the 3' splice site is an intermolecular version of exon ligation, then each precursor RNA should show the greatest reactivity with the dinucleotide whose sequence is homologous to that of the 3' end of its 5' exon. This prediction is satisfied by the relative reactivity of the *Neurospora* pre-mRNA with pGpU compared with other dinucleotides and of the *Tetrahymena* pre-rRNA with pCpU compared with other dinucleotides.

The relative reactivities of the different dinucleotides is most easily explained if each intron contains an exon-binding site that is complementary to the last two nucleotides of its own 5' exon. This requirement is fully met by the internal guide sequence model of Davies *et al.*². The internal guide has a different sequence in each intron, but a portion of the sequence is always complementary to the 6 or so nucleotides in the 5' exon immediately preceding the intron. The only conserved nucleotide is a G in the internal guide which pairs with the conserved U at the 5' splice site. Similar structural models incorporating the 5' exon-internal guide sequence pairing have been proposed by Michel *et al.*^{4,5} and Bos *et al.*¹¹.

Figure 3 shows the proposed mechanism by which the internal guide sequence promotes the intermolecular exon ligation reactions. The 5' exon is bound to the internal guide, but in the absence of cleavage by guanosine the exon does not have a free 3' hydroxyl group, so it cannot attack the 3' splice site. When the dinucleotide transiently displaces the 5' exon, its free 3' hydroxyl group attacks the 3' splice site.

In the normal exon ligation reaction, the 5' exon-binding site is proposed to serve two functions. It is responsible for choice of the 5' splice site, that is, the site at which free guanosine or GTP attacks. Thus, the guanosine-binding site¹⁹ must be in a fixed position relative to the 5' exon-binding site, perhaps by way of a direct interaction⁹. In addition, the binding site orients the 5' exon in proximity to the 3' splice site, helping to lower the activation energy for the exon ligation reaction. It therefore helps to determine the accuracy of reactions at both the 5' and 3' splice sites.

Recently, site-specific mutagenesis studies have confirmed that the sequence GGAGGG at nucleotides 22–27 of the *Tetrahymena* intron is the 5' exon-binding site^{26,27}.

We thank S. Walstrum and O. Uhlenbeck (University of Colorado) for their gift of trinucleotides. This work was supported by NIH grant GM26836 to A.M.L. and by NIH grant GM28039 and American Cancer Society grant NP374B to T.R.C.

Received 17 March; accepted 15 April 1986.

- Burke, J. & RajBhandary, U. L. *Cell* **31**, 509–520 (1982).
- Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A. & Scazzocchio, C. *Nature* **300**, 719–724 (1982).
- Waring, R. B., Scazzocchio, C., Brown, T. A. & Davies, R. W. *J. molec. Biol.* **167**, 595–605 (1983).
- Michel, F., Jacquier, A. & Dujon, B. *Biochimie* **64**, 867–881 (1982).
- Michel, F. & Dujon, B. *EMBO J.* **2**, 33–38 (1983).
- Cech, T. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 3903–3907 (1983).
- Zaug, A. J., Grabowski, P. J. & Cech, T. R. *Nature* **301**, 578–583 (1983).
- Cech, T. R. *Cell* **44**, 207–210 (1986).
- Inoue, T., Sullivan, F. X. & Cech, T. R. *J. molec. biol.* **189**, 143–165 (1986).
- Inoue, T., Sullivan, F. X. & Cech, T. R. *Cell* **43**, 431–437 (1985).
- Bos, J. L. *et al. Cell* **20**, 207–214 (1980).
- Garriga, G. & Lambowitz, A. M. *Cell* **39**, 631–641 (1984).
- Tabak, H. F., Van der Horst, G., Osinga, K. A. & Arnberg, A. C. *Cell* **39**, 623–629 (1984).
- Hengstens, L. A. M., Bonen, L., de Haan, M., Van der Horst, G. & Grivell, L. A. *Cell* **32**, 379–389 (1983).
- Van der Horst, G. & Tabak, H. F. *Cell* **40**, 759–766 (1985).
- Sullivan, F. X. & Cech, T. R. *Cell* **42**, 639–648 (1985).
- Helmer-Citterich, M., Morelli, G. & Macino, G. *EMBO J.* **2**, 1235–1242 (1983).
- Burke, J. M., Breitenberger, C., Heckman, J. E., Dujon, B. & RajBhandary, U. L. *J. biol. Chem.* **259**, 504–511 (1984).
- Bass, B. L. & Cech, T. R. *Nature* **308**, 820–826 (1984).
- Price, J. V. & Cech, T. R. *Science* **228**, 719–722 (1985).
- McMaster, G. K. & Carmichael, G. G. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4835–4838 (1977).
- Donis-Keller, H. *Nucleic Acids Res.* **7**, 179–192 (1979).
- Lockard, R. E. *et al. Nucleic Acids Res.* **5**, 37–56 (1978).
- Donis-Keller, H. *Nucleic Acids Res.* **8**, 3133–3142 (1980).
- Been, M. D. & Cech, T. R. *Nucleic Acids Res.* **13**, 8389–8408 (1985).
- Waring, R. B., Towner, P., Minter, S. J. & Davies, R. W. *Nature* **321**, 133–139 (1986).
- Been, M. D. & Cech, T. R. (submitted).

Intermolecular binding of xanthan gum and carob gum

P. Cairns, M. J. Miles and V. J. Morris

AFRC Institute of Food Research, Norwich Laboratory, Colney Lane, Norwich NR4 7UA, UK

Gels are a representative state for polysaccharides in both natural and artificial systems. The nature of the inter-chain associations within the junction zones is important and models for such interactions between like polysaccharides are based on X-ray diffraction studies of oriented gels. Here we describe the extension of such studies to a binary gel (xanthan–carob) in order to characterize for the first time intermolecular binding between different polysaccharides. Xanthan–carob binding has been proposed to explain gelation of the mixtures and as a model for host–pathogen recognition and adhesion of *Xanthomonas* bacteria within plant vascular systems. Our data suggest that the established model^{1–7} is incorrect and point to an alternative association mechanism.

Xanthan has a repeat unit^{8,9} based on a cellulose backbone with alternate glucose residues O-3-substituted with a trisaccharide side chain. Carob has a mannan backbone that is incompletely O-6-substituted with galactose² and a mannose/galactose ratio of 3.55 (ref. 10). Xanthan–carob binding is postulated¹ to involve a mixed cooperative interaction in which the xanthan helix is retained (Fig. 1). We tested this model by X-ray fibre diffraction studies of oriented xanthan–carob gels, since this is

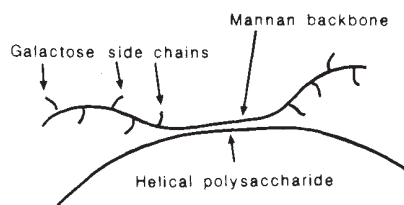


Fig. 1 Schematic diagram of the proposed¹ binding of the xanthan helix to the bare mannan backbone of the galactomannan carob. Analogous models have been proposed^{1–4} to explain the interaction between galactomannans and certain algal polysaccharides.

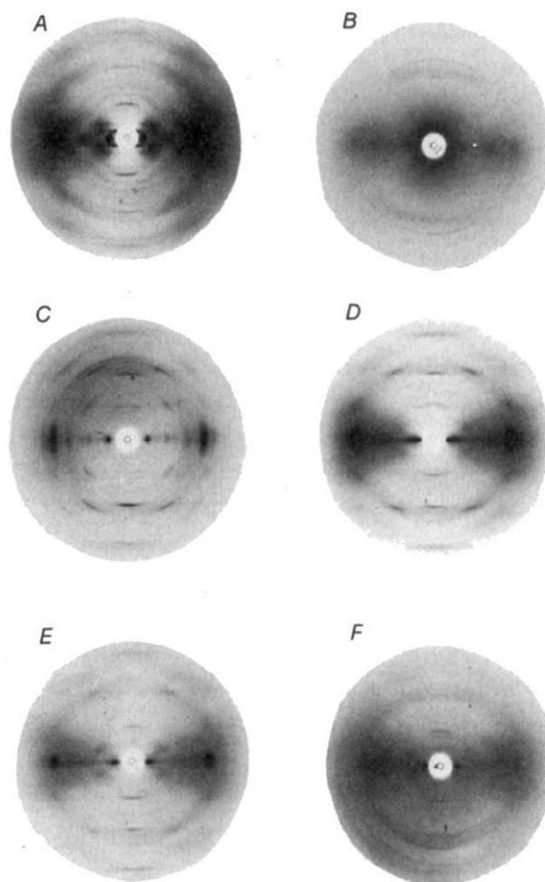


Fig. 2 X-ray fibre diffraction patterns recorded for polysaccharides and polysaccharide mixtures. Gels were cut into strips and stretched to induce alignment. Non-gelling mixtures were cast onto glass or PTFE substrates, dried to films, cut into strips and aligned by stretching. The X-ray wavelength was 0.154 nm; the camera was flushed with helium to reduce scattering. Pictures were calibrated with calcite. A, Xanthan, stretched 300%; B, carob, stretched 300%; C, carob, stretched 300% and stored at room temperature for 3 yr; D, mixed gel: 25% xanthan/75% carob, stretched 300%; E, mixed gel: 50% xanthan/50% carob, stretched 300%; F, non-gelling mixture: 50% xanthan/50% carob, prepared at room temperature, stretched 100%.

the only method available to examine polysaccharide ordered structures at atomic resolution and has provided reliable models for structures in the hydrated state¹¹.

Under conditions appropriate for studying fibres prepared from mixed gels, pure xanthan and carob gave the fibre patterns shown in Fig. 2A, B. Storage of carob fibres enhanced crystallinity (Fig. 2C). Diffraction data obtained for fibres prepared from xanthan–carob mixed gels are shown in Fig. 2D, E and these new X-ray patterns provide evidence for xanthan–carob binding.

Optical rotation measurements are sensitive to the helix–coil transition¹² and were used to monitor helix formation in xanthan samples. Thorough mixing of carob with xanthan in the helical conformation at room temperature did not lead to gelation. X-ray diffraction data on fibres prepared from such mixtures (Fig. 2F) showed reflections characteristic of pure xanthan (Fig. 2A) with no evidence of carob crystallization or carob–xanthan co-crystallization. Gelation only occurred when these mixtures were heated (95 °C) above the xanthan helix–coil transition temperature (T_c) and then re-cooled to room temperature. Oriented gels yielded fibre patterns of the type shown in Fig. 2D, E. T_c is particularly sensitive to the presence of divalent cations

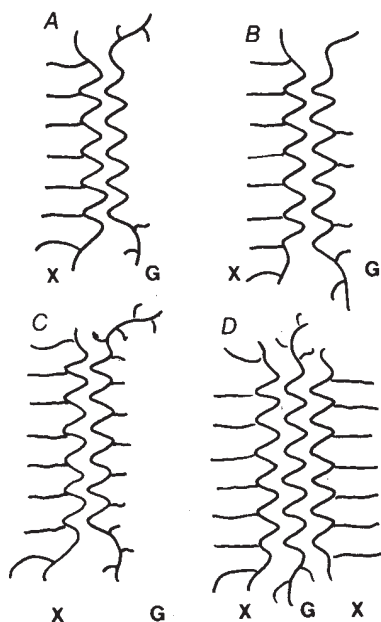


Fig. 3 Schematic representation of possible xanthan (X)-galactomannan (G) binding. Possible binding between a xanthan backbone and: *A*, bare mannan regions of the galactomannan backbone; *B*, randomly substituted galactomannan; *C*, galactomannan containing galactose on alternate mannose residues. *D*, Simplest xanthan-galactomannan sandwich structure.

in the medium¹³. When carob was mixed with xanthan in the presence of sufficient calcium chloride such that $T_c > 100^\circ\text{C}$, and the samples were heated (95°C) and re-cooled to room temperature, they did not gel; this suggests that denaturation of the xanthan helix is necessary if intermolecular association and gelation is to occur. Glucose and mannose differ only in the orientation of the hydroxyl substituent at C-2, suggesting the possibility of binding between the sterically compatible cellulose backbone of xanthan and the mannan backbone of carob.

Optical rotation studies¹⁴ on xanthan-galactomannan mixtures coupled with enzymatic studies¹⁵ on xanthan-carob mixtures suggest that only small segments of both backbones could be involved in binding, with the remaining xanthan reforming the helical conformation. Alignment of mixed junction zones will also lead to alignment of xanthan helices. Thus, reflections characteristic of aligned xanthan should be neglected in Fig. 2*D*, *E* in order to characterize the mixed junction zone. The first meridional reflection for the mixed junction zones corresponds to an interplanar spacing of 0.52 nm, characteristic of cellulose¹⁶ or mannan¹⁷, supporting binding between the backbones but ruling out simple binding schemes of the type shown in Fig. 3*A-C*. To achieve the correct interplanar spacing appropriate to the first meridional reflection, it would be necessary to stagger the positions of the xanthan side chains (Fig. 3*D*) although the stoichiometry of such sandwich structures is undefined. The mixed junction zone patterns are equivalent to carob patterns for which only *0kl* reflections are allowed. The simplest answer would be to permit growth of a sandwich structure in the *b* and *c* directions but to restrict growth in the *a* direction. However, streaking of layer line reflections characteristic of such laminar structures was not observed. The alternative is to envisage a structure periodic in the *b* and *c* directions but aperiodic in the *a* direction.

The data presented here constitute the first example of a conformational modification of one polysaccharide by a non-covalent interaction with another polysaccharide. The structural similarity of the xanthan and carob backbones accounts for the

specificity of the interaction and also explains why X-ray diffraction studies^{10,18,19} have failed to reveal proposed¹⁻⁷ intermolecular binding between galactomannans and the structurally incompatible algal polysaccharides κ -carrageenan and furcellaran. The data support suggested co-crystallization of cellulose with galactomannans²⁰ or xyloglucans²¹⁻²³ in complex plant cell walls. Binding of extracellular xanthan to plant cell wall components is only likely if helix formation is incomplete. This cannot, however, exclude possible host-pathogen interactions because such extracellular assembly is normally biochemically controlled.

Received 17 March; accepted 9 May 1986.

1. Morris, E. R., Rees, D. A., Young, G., Walkinshaw, M. D. & Darke, A. *J. molec. Biol.* **110**, 1-16 (1977).
2. Dea, I. C. M. & Morrison, A. *Adv. Carbohydr. Chem. Biochem.* **31**, 241-312 (1975).
3. Dea, I. C. M. *Am. chem. Soc. Symp. Ser.* **150**, 439-454 (1981).
4. Dea, I. C. M., McKinnon, A. A. & Rees, D. A. *J. molec. Biol.* **68**, 153-172 (1972).
5. McCleary, B. V. *Carbohydr. Res.* **71**, 205-230 (1979).
6. Tako, M., Asato, A. & Nakamura, S. *Agric. biol. Chem.* **48**, 2995-3000 (1984).
7. Tako, M. & Nakamura, S. *Carbohydr. Res.* **138**, 207-213 (1985).
8. Jansson, P. E., Kenne, L. & Linberg, B. *Carbohydr. Res.* **45**, 275-282 (1975).
9. Melton, L. D., Mindt, L., Rees, D. A. & Sanderson, G. R. *Carbohydr. Res.* **46**, 245-257 (1976).
10. Miles, M. J., Morris, V. J. & Carroll, V. *Macromolecules* **17**, 2443-2445 (1984).
11. Rees, D. A., Morris, E. R., Thom, D. & Madden, J. R. in *The Polysaccharides* Vol. 1 (ed. Aspinall, G. O.) 195-290 (Academic, London, 1982).
12. Norton, I. T., Goodall, D. M., Frangou, S. A., Morris, E. R. & Rees, D. A. *J. molec. Biol.* **175**, 371-394 (1984).
13. Lambert, F., Milas, M. & Rinaudo, M. *Int. J. biol. Macromolec.* **7**, 49-52 (1985).
14. Dea, I. C. M. *et al. Carbohydr. Res.* **57**, 249-272 (1977).
15. McCleary, B. V., Amado, R., Waibel, R. & Neukom, H. *Carbohydr. Res.* **92**, 269-285 (1981).
16. Blackwell, J., Gardner, K. H., Kolpak, F. J., Minke, R. & Claffey, W. B. *Am. chem. Soc. Symp. Ser.* **141**, 315-334 (1980).
17. Frei, E. & Preston, R. D. *Proc. R. Soc. B* **169**, 127-145 (1968).
18. Cairns, P., Miles, M. J. & Morris, V. J. *Int. J. biol. Macromolec.* **8**, 124-127 (1986).
19. Cairns, P., Miles, M. J., Morris, V. J. & Brownsey, G. J. *Carbohydr. Res.* (in the press).
20. Marchessault, R. H., Buléon, A., Deslandes, Y. & Goto, T. *J. Colloid Interface Sci.* **71**, 375-382 (1979).
21. Bauer, W. D., Talmadge, K. W., Keegstra, K. & Albersheim, P. *Pl. Physiol.* **54**, 174-187 (1973).
22. Aspinall, G. O., Molloy, J. A. & Graig, J. W. T. *Can. J. Biochem.* **47**, 1063-1070 (1969).
23. Taylor, I. E. P. & Atkins, E. D. T. *FEBS Lett.* **181**, 300-302 (1985).

Erratum

Molecular distinction between fetal and adult forms of muscle acetylcholine receptor

M. Mishina, T. Takai, K. Imoto, M. Noda, T. Takahashi, S. Numa, C. Methfessel & B. Sakmann

Nature **321**, 406-411 (1986)

THE first sentence in the summary to this article should read: "Distinct classes of acetylcholine receptor channels are injected with combinations of the bovine α -, β -, γ - and δ - or the α -, β -, δ - and ϵ -subunit-specific messenger RNAs."

Corrigendum

Near-ultraviolet and visible spectrophotometry of comet Halley from Vega 2

G. Moreels, M. Gogoshev, V. A. Krasnopolsky, J. Clairemidi, M. Vincent, J. P. Parisot, J. L. Bertaux, J. E. Blamont, M. C. Festou, Ts. Gogosheva, S. Sargoichev, K. Palasov, V. I. Moroz, A. A. Krysko & V. Vanysek

Nature **321**, 271-273 (1986)

IN this letter in *Nature's* comet Halley supplement, an incorrect address was given for one of the authors. V. Vanysek's correct address is: Charles University, 15000 Prague, Czechoslovakia.

Imaging cells with acoustic microscopy

from M. Issouckis

The interaction between the mechanical waves of scanning acoustic microscopy (SAM) and the mechanical features of living cells makes a novel contribution to the characterization of biological materials.

HIGH resolution scanning acoustic microscopy, or SAM, provides an opportunity to investigate the microstructure of materials on, and under, the material surface. In biology, acoustic imaging of living cells, mucous coats and cuticle structures has been very successful. The imaging of these biological objects has been described by M. Hoppe and J. Bereiter-Hahn¹. This paper uses their work with living cells¹ as a springboard for discussing high resolution SAM and SAM image interpretation in the life sciences.

The heart of the high resolution SAM (Fig.1) is its lens assembly. A focused ultrasound wave (of frequencies between 1 and 2 GHz), generated by means of a piezoelectric transducer, is used to probe the sample. This same transducer receives the reflected wave, which is subsequently processed. By scanning in a raster fashion, an image is formed that represents the interaction between the acoustic material and the elastic properties of the sample material.

Some requirements of SAM imaging impose limitations on the types of samples that can be studied. As in high resolution light microscopy, high resolution SAM requires lenses with short focal length. Consequently, the depth of focus is small. In addition, attenuation (absorption) of acoustic waves by the intermedium limits the distance between lens and specimen.

These features restrict the application of acoustic microscopy to very small objects; that is, those that are very flat and in a plane parallel to the scanning plane of the acoustic lens. Flat structures such as insect wings and scales, and cells attached to glass or any other planar surface, fulfil these conditions.

Although observations of cell or organelle damage by ultrasound have been reported for lower frequencies (20 kHz, 5 MHz)^{2,3} damage to living cells does not occur with gigahertz frequencies. Experiments have shown that after more than one hour's exposure to 1 GHz ultrasound, sensitive epidermal cells still exhibit normal locomotive behaviour. These findings are consistent with theoretical expectations, which predict a decrease in cell damage as frequency increases because the amplitude of particle displacement induced by a longitudinal wave is limited by the wavelength.

Interpretation

Some properties of SAM can also confound image interpretation. Images of biological samples, in particular, often

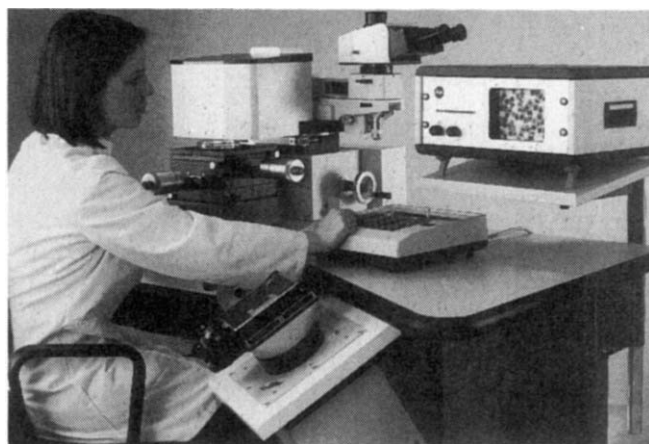
show interference lines due to the substructure supporting the sample. In materials science applications, the smooth surface of a specimen represents the structure that is to be investigated. In the case of biological objects, however, the samples are located on a planar surface to which they adhere.

This situation complicates the interpretation of acoustic images by producing

on the incompressible interior cytoplasm^{8,9}.

The mechanical properties and surface profiles of living cells result from the interaction of elements of the cytoskeleton with each other, with the cell membrane, and with the supporting substratum. Microtubules and stress fibres stiffen the cytoplasm and, together with adhesion to a substratum, cause deviations of cell shape from a sphere. This model allows an

Fig.1 An Ernst Leitz scanning acoustic microscope (ELSAM) in its commercial configuration.



interference between the waves reflected at different interfaces. These interfaces include the surface of the substratum, the medium-facing surface of the cells and, depending on the distance of a cell to the substratum, the substratum-facing cell surface.

Comparable relations are found in reflection-contrast microscopy⁴⁻⁷. Light reflected from separate interfaces causes a pattern of interference lines. The pattern indicates the surface curvature of cell at the medium-facing surface and at the surface opposed to the substrate. In this case, optical path differences that determine the course of the interference lines depend on refractive indexes and the thickness of the layers between interfaces.

Similarly, with the scanning acoustic microscope, layer thickness and acoustic impedance (which corresponds to the refractive index in optics) govern the interference pattern. But whereas light absorption is negligible in most living cells, attenuation can contribute considerably to image formation in SAM.

The internal structures of cells can also alter image generation by reflection and scattering. In general, an animal cell can be considered as a hydraulic system enclosed by a semipermeable membrane to which a contractile fibrillar system (containing actin and probably myosin) is linked. This system exerts some pressure

understanding of acoustic images of living cells.

Mechanical disturbances by the scanning movement of the SAM lens and flow of the interconnecting medium, which are caused by sound emission in the cavity of the lens, loosen the contact of cells to a substratum. In those cells that withstand this stress, density in stress fibre arrangement is enhanced considerably.

Internal reflections

In the images of *Xenopus* heart cells (Fig. 2), stress fibres are clearly shown as arcs near the periphery of the cell. The dark circles concentric with the nucleus. The difference in contrast between the sample mounted on plastic (top panel) and the sample mounted on glass (bottom panel) is apparent.

The peripheral region of the plastic-mounted cell is relatively flat; its thickness is approximately $\lambda/2$ at the outer dark interference ring, increasing to about $\lambda/4$ in the outer bright region. For these estimations a phase shift of $\lambda/2$ at the interface to the medium with higher acoustic impedance is considered.

The small distances between central interferences indicate an increased slope; that is, the thickness of the sample increases in the region. The contrast of an interference line may vary considerably throughout its course, so that a dark line may

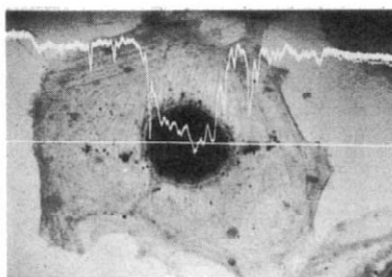
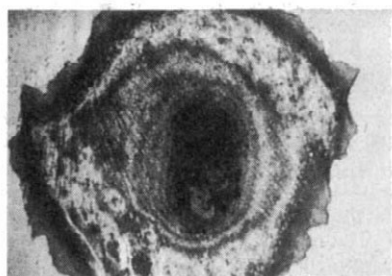


Fig. 2 SAM images of endothelial cells from *Xenopus laevis* tadpole hearts (XTH-2 cells) using $\times 625$ magnification and a frequency of 1.6 GHz. The image width is 200 μm . **Top panel:** a sample mounted on plastic gives an interference image with stress fibres and cell topography clearly visible. The contrast of the cell fringes is related to stiffness conferred by intracellular structures. **Bottom panel:** a sample mounted on glass gives a more poorly defined attenuation image.

become very faint at some points (Fig. 2). However, this is not caused by a local thickness change, but by locally diminished reflectivity.

The structural basis for this reduced reflectivity is a weakened stiffness, visualized by the low contrast that is due to a diminished elastic modulus of the cortical cytoplasm. At this "weak region" the membrane can bulge out, and a thickening appears. The dark interference fringe close to the cell margin increases in contrast, indicating higher reflectivity.

The complexities posed by interference and contrast changes make SAM image interpretation a challenging task, but the information that can be gleaned from this complexity makes scanning acoustic microscopy a valuable technique for exploring the substructure of the smallest unit of life.

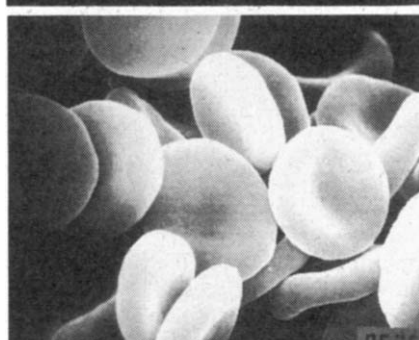
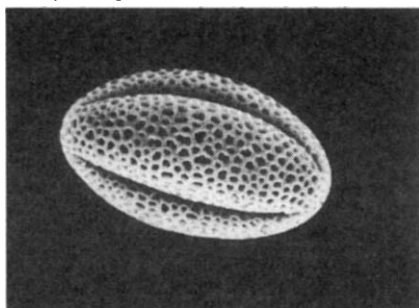
Matthew Issouckis is at E. Leitz Instruments Ltd, 48 Park Street, Luton LU1 2HP, UK.

- Hoppe, M. & Bereiter-Hahn, J. *IEEE Trans. Sonics Ultrasonics* Vol. SU-32 (2), 289 (1985).
- Cachon, J. & Chachon, M. *Biol. Cell*, **40**, 23-32 (1981).
- Cachon, J., Cachon, M. & Bruneton, J.N. *Biol. Cell* **40**, 69-72 (1981).
- Bereiter-Hahn, J., Fox, C.H. & Thorell, B. *J. cell. Biol.* **82**, 767-779 (1979).
- Beck, K. & Bereiter-Hahn, J. *Microchim. Acta*, **84**, 153-178 (1981).
- Gingell, D. *J. cell. Sci.* **49**, 237-247 (1981).
- Gingell, D. & Todd, I. *J. biophys. biochem. Cytol.* **26**, 507-526 (1979).
- Bereiter-Hahn, J., Strohmeier, R., Kunzenbacher, I., Beck, K. & Voth, M. *J. cell. Sci.* **52**, 289-311 (1981).
- Strohmeier, R. & Bereiter-Hahn, J. *Expl cell. Res.* **154**, 412-420 (1984).

Setting the stage for microscopy

A collection of instruments and accessories for microscopy and image analysis closes with immunology product selections.

JEOL is offering a **scanning electron microscope** at a price that is more typical of advanced optical models (Reader Service No. 100). The T100, a compact and relatively simple instrument, is aimed at



Micrographs from Jeol's new 'scopes. Rape pollen, top panel; control blood, bottom panel.

laboratories engaged in medical or botanical research and sold for £15,000 (UK). Image contrast and brightness controls are available along with a fast scan facility for TV viewing. The T100 features a single magnification adjustment that spans $\times 15$ to $\times 100,000$ with electronic readout. Equipped with several information channels, Jeol's microscope can be fitted with attachments for elemental analyses such as characteristic X-rays and cathodoluminescence.

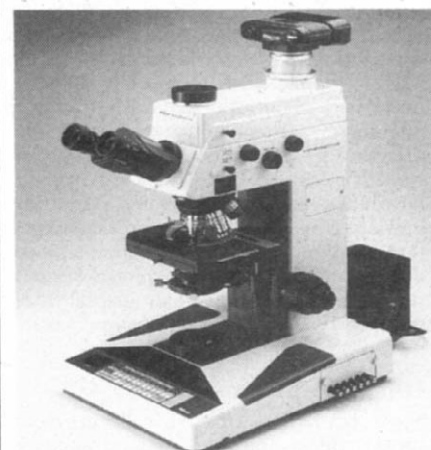
A **photomicrographic microscope system** for research goes under the name of the New Vanox S/T from Olympus (Reader Service No. 101). Olympus says it has provided the Vanox with a flexibility that allows a wide variety of uses. The instrument has LB series objectives for super widefield observation; the eyepiece field number reaches 26.5. A motorized revolving nosepiece will alternate between six objectives at the touch of a button, whilst Köhler illumination shifts the lighting for the range of magnifications between ultra-low and high magnifying objectives. Four photographic eyepieces



Olympus makes flexibility its forte.

supply magnifications between $2.5\times$ and $5\times$. Olympus designed its £18,000 (UK) machine with a rotatable stage for proper specimen alignment, and a light intensity control system based on ND filter conversion maintains constant colour temperatures regardless of light intensity.

The Microphot-FX, Nikon's new **photomicrography system**, combines the optical performance of the CF Plan Apocromat and Achromat series with versatile photomicrography (Reader Service No. 102). Nikon says the new instrument has 1 per cent spot or 30 per cent average exposure measurement switching, direct light measurement, direct image projection, and movable measurement area. The Microphot is also available with the



The Microphot design for photomicrography.

photographic attachment as an option, when photomicrography is a secondary purpose. Nikon also make several optical accessories as options, including phase contrast, diascopic fluorescence and interferometry equipment.

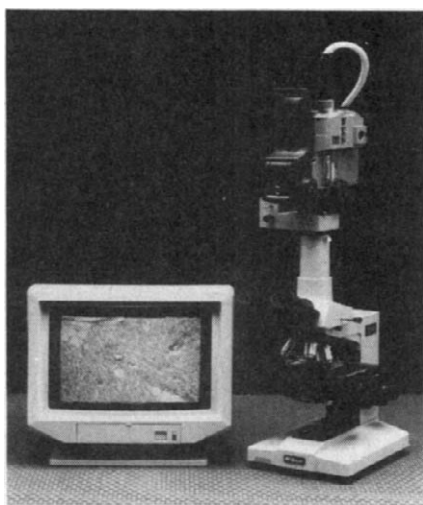
Straight from West Germany, an expanded range of Zeiss microscopes is being distributed by Gallenkamp in the United Kingdom (Reader Service No. 103). The company now includes "budget priced" **stereo and compound microscopes**, as well as more sophisticated instruments from the manufacturer. Microscopes for student work and routine research applications have been added to



Zeiss technology in the UK via Gallenkamp.

the list. Gallenkamp claims its new selection is ideal for biological and materials examination.

Joining Philips' display at Micro '86 will be the company's latest **scanning electron**



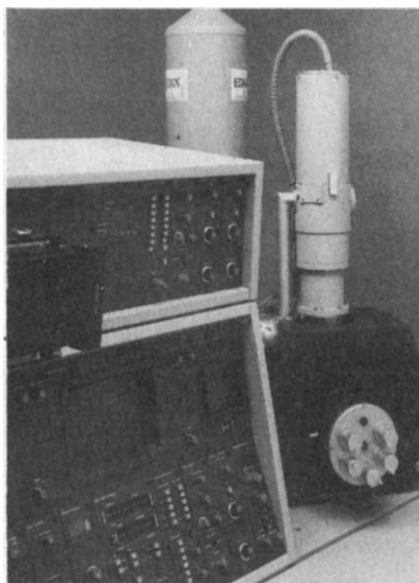
Part of Philips' Micro '86 display.

microscope, the 525M (Reader Service No. 104). The SEM 525M, part of a field of multi-purpose microscopes, offers a large specimen chamber and motorized stage with six-inch movement. Philips will

also have on exhibit the CM10, a microprocessor-controlled transmission electron microscope. The system operates through a small VDU with a divided screen and soft-key controls. An energy-dispersive analyser set-up called EDAX PV 9900 will also make an appearance at the conference.

A sharper image

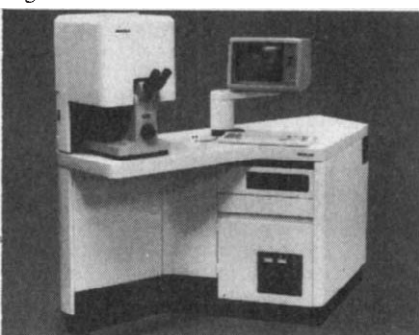
For full **colour image analysis**, Analytical Measuring Systems has a semi-automatic



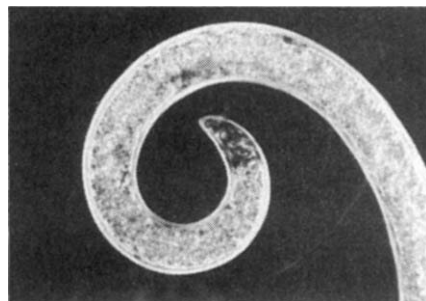
AMS has analysis in full colour.

system with software that runs on IBM PCs or compatibles (Reader Service No. 105). The VIDS III Colour uses a three-gun video camera to obtain micro- or macroscopic sample images, and according to AMS, is particularly well-suited to applications in histology, cytochemistry and mineralogy. The VIDS III software can measure features within the images that are designated using a graphics tablet.

A **fluorescence analysis workstation** from Meridian Instruments can perform quantitative analyses of anchorage-dependent cells without disturbing sterile conditions (Reader Service No. 106). Meridian's ACAS 470 system employs computer control and x-y stage positioning to direct a focused laser beam. Whilst



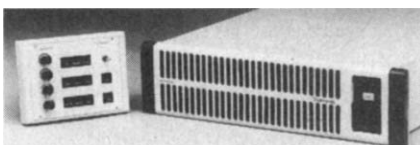
The flow cytometer for stationary cells.



A spiral nematode posed for this photomicrograph, which won the Nikon Small World competition in 1985. (Reader Service No. 107). Nikon's annual contest is international and open to anyone interested in photography through microscopes. Both subject matter and microscopy technique are unrestricted. Past winners have come from disciplines as diverse as marine science and dentistry. This year's deadline was 27 June, but it is not too early to be eyeing up the 1987 competition. Twenty contestants will receive awards, with the first place photographer winning a \$3,000 vacation or the equivalent in Nikon equipment. Last year's winner, who captured the image above, was J.D. Eisenback from Blacksburg, Virginia.

preserving cell attachment to extracellular matrix, the beam carries out optimal fluorescence excitation, cell sorting by laser ablation or physical separation on film, cell surgery and photobleaching of discrete areas. The workstation can also perform automated repetitive analysis on single cells or fields of cells.

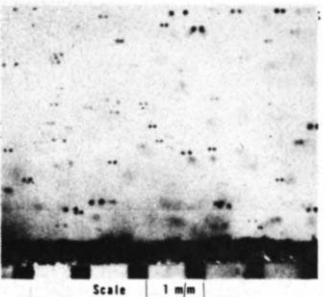
Quantel is exhibiting a high-speed **digital image processing device** at Micro '86 that it says can improve the contrast and



Making images CRYSTAL clear.

sharpness of images to analytical quality in real-time (Reader Service No. 108). Quantel calls its system CRYSTAL and claims the instrument can increase through-put as well as statistical validity. CRYSTAL also includes a second frame-store for a reference image that can be compared to the sample input image in real-time. Quantel says CRYSTAL's features are beneficial in the inspection and analysis of non-conducting specimens. The £15,500 (UK) system is effective with ion beam, scanning optical and scanning microscopes as well as electron microscopes.

Polyscan is a fully **automatic image analyser** with flexible software control (Reader Service No. 109). Reichert-Jung



SPRAY ANALYSIS

PULSE

'PULSE' instrumentation equipment will produce time separated images of sprays and aerosols as illustrated in this photograph. If you are concerned with the atomization of FUELS - SOLVENTS - PAINTS - CHEMICALS this picture is important to you. VELOCITY - SIZING - DISTRIBUTION can be assessed from such pictures which will help you improve the efficiency of your spraying equipment and conserve expensive materials.

Reader Service No.14

SINGER INSTRUMENTS

Xenopus oocytes injection?



SINGER MKI MICROMANIPULATOR

We also offer a Microsyringe Kit.

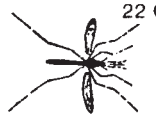
Write to:
Singer Instrument Co. Ltd.
Treborough Lodge
Roadwater, Watchet
Somerset TA23 0QL
England. Tel. (0984) 40226
Telex: 337492 Comcab G

Reader Service No.50

Activity Monitor

Track that rat, insect, fish, limb

Tracking unit accepts input from video camera or recorder, and outputs the location of a high contrast object in the TV image. Interfaces for BBC, Apple, IBM-PC.



HVS IMAGE ANALYSING LTD,
22 Cromwell Road, Kingston,
KT2 6RE, England.
Tel: 01-546 9090

Reader Service No.16

Medwire Div.
stocks a wide variety of

TEFLON COATED*

Platinum alloys, stainless steel, silver, multistrand and single strand available for prompt shipment. Write for descriptive brochure and price lists.



Medwire Div.
Affiliate of Sigmund Cohn Corp.
123 So. Columbus Ave.
Mt. Vernon, N.Y. 10553
914-664-5300
*Reg. T.M. of E.I. DuPont

Reader Service No.49

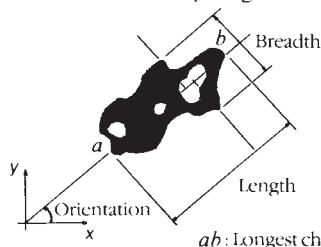


Processing by Polyscan can be customized.

developed the system with an extensive image processing library that enables customized programs to be constructed using a program editor. Polyscan accepts inputs from microscopes, video tape recorders, photometers and photographs and uses synchronous network architecture for real-time processing. For complex images, the system also allows semi-automatic measurement via a cursor-controlled digitizer.

Digithurst Ltd makes non-contact measurement and **image analysis packages** that perform data processing with standard microcomputers such as the IBM PC (Reader Service No. 110). The company's range can accommodate most types of microscopes, with camera mounts for optical microscopy. Digithurst's Micro-Scale software directs the detection of cell boundaries and outlines, growth rate determinations and measurements including area, length, counting, orientation to field and minimum and maximum diameters. A MicroStat statistics package provides derivation of statistical results and graphical presentation.

Joyce-Loebl introduced an **image analysis system** earlier this year called μ Magiscan that will be displayed at Micro '86 (Reader Service No. 111). The package is based on a modified IBM PC, with software composed of test, analysis and statistics routines. The μ Magiscan analysis



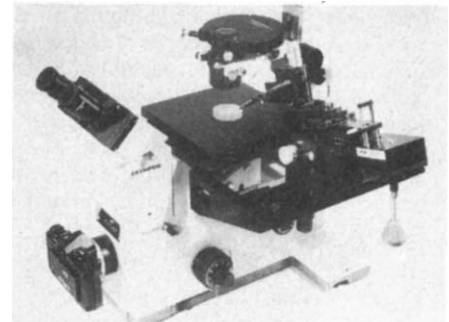
Joyce-Loebl's analyser takes on new dimensions.

menu is capable of establishing working conditions, pre-processing to optimize the image state and detecting, selecting and measuring designated objects. Joyce-Loebl combined several hardware features to create μ Magiscan, including a colour graphics adaptor board, a 10 MB Winchester disk and a Microsoft Mouse. The total μ Magiscan package sells for less than £15,000 (UK) and consists of the computer, graphics boards, colour moni-

tor, camera, mouse, software and an Epson printer.

Microtools

Inverted microscopes such as the Olympus IMT-2, the Nikon Diaphot and the Zeiss IM35 are the targets for a **micromanipulator** from Research Instruments Ltd (Reader Service No. 112). RI claims



A micromanipulator that specializes in inversion.

that the TDU500 is the first lever-operated micromanipulator to use flexural hinges in its fine movement. According to the company, these hinges can provide sub-micrometre positioning accuracy. The instrument has a single lever for x - y - z fine movement, and levels for coarse x , y and z motion. Toolholders for one or two microtools are available on different TDU500 models.

Landmarks aren't easy to find under magnification, so EBETC Corp sells a \$63 (US) **slide marker** that puts an inked ring on slide locales that need to be identified (Reader Service No. 113). The ring is



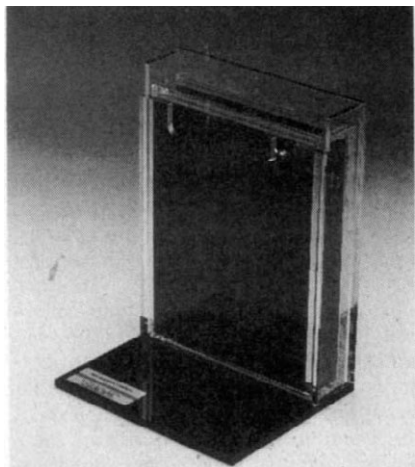
EBTEC marks the spot on microscope slides.

repeatable and concentric with the optical axis; marking will not break or damage slides. EBTEC's device fits any standard optical microscope with RMS objective thread, and the mark can be removed with solvent. EBTEC calls the marker "indispensable" for cytology, pathology, electron microscopy and industry applications.

Immunology addendum

Using plate electrodes to generate a uniform field, an **electrophoretic blotting apparatus** from Idea Scientific will execute Western blots in 30 minutes (Reader Service No. 114). Idea's GENIE has titanium

plates coated with a thin layer of platinum which can blot gels up to 17 cm long. The GENIE also features an insertable loading tray that the company says will help achieve bubble-free loading. Idea's



The uniform field theory comes to electrophoresis.

machine requires less than one litre of buffer. Another version is available for gels up to 32 cm long.

Becton-Dickinson has recently released a new **monoclonal antibody for B cell studies** (Reader Service No. 115). The antibody, called Anti-Leu-16, reacts with an antigen found on human B cells in peripheral blood, tonsil and lymph node tis-



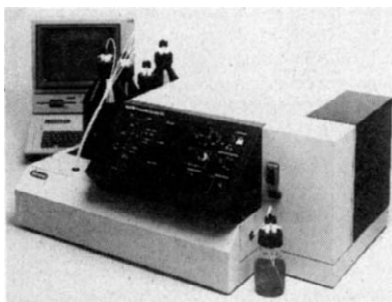
Becton-Dickinson's tool for trailing B cells.

sue. The antigen appears on approximately 10 per cent of peripheral blood lymphocytes, according to B-D. The company also says Anti-Leu-16 will not react with T cells or granulocytes and exhibits marginal reaction with monocytes. Available in purified form or as a fluorescein or phycoerythrin conjugate, the antibody can be used for B cell enumeration in peripheral blood, B cell isolation and studies of B cell neoplasms and activation.

Erratum: streptavidin pricing

A review published in the 29 May issue of *Nature* (Vol. 321, p. 543) gave an incorrect price for Porton Products Ltd's streptavidin. In the United Kingdom, the chemical costs £5 per 0.5 mg, not 100 mg as previously stated. Porton Products sells streptavidin in 100 mg quantities for £1,000.

The latest version of Bio-Rad's MAPS (Monoclonal Antibody Purification System) kit has been designed specifically



MAPS dedicated to particular purifications.

for the **purification of mouse IgG** from ascites fluid (Reader Service No. 116). The kit, called Affi-Gel Protein A MAPS II, consists of protein A with MAPS buffers optimized for binding, elution and regeneration. Bio-Rad says the binding buffer increases the protein A-agarose capacity for IgG from less than 1 mg per ml to 6 to 8 mg per ml of gel. The elution buffer, the company claims, gives quantitative elution of bound IgG, and the regeneration buffer provides for up to 12 cycles. The entire kit contains the materials necessary to purify 500 mg of murine antibody.



Applications for Skatron's new harvester keep growing.

Skatron is promoting a new **cell harvester** as an aid in receptor binding studies (Reader Service No. 117). The company describes a procedure that is intended to reduce the time and costs involved in filtration and reduce the levels of non-specific binding. With Skatron's instrument, twelve incubation samples can be simultaneously filtered under reduced pressure. Another component, the Filter-punch, transfers the filter disks six at a time to scintillation vials. Skatron's cell harvester can also be employed in studies of HTLV-III/LAV infections. □

These notes are compiled by Karen Wright from information provided by the manufacturers. To obtain further details about these products, use the reader service card bound inside the journal.

Prices quoted are sometimes nominal and apply only within the country indicated.

Roland GRUNBERG - NANCY - FRANCE

DO
YOU
NEED
PEPTIDES?

NEOSYSTEM
LABORATORIES

PROVIDE
A COMPLETE SERVICE FOR
PEPTIDE SYNTHESIS
PEPTIDE AND PROTEIN SEQUENCING
PEPTIDE ANTIBODY PRODUCTION

WE PROVIDE ALSO
CONTRACT RESEARCH SERVICES
TO INDUSTRIAL AND ACADEMIC
INSTITUTIONS



NEOSYSTEM
LABORATORIES

B.P. n° 309 R8
67008 STRASBOURG Cedex
FRANCE - Télex 880 088 F
Tél. : 88.36.96.96

Reader Service No.28

Will industry support higher education?

from Richard Pearson

Since 1980, government support for UK universities has dropped by over 11 per cent in real terms. Can new initiatives from industry soften the blow?

IN the United Kingdom, higher education is now halfway through a decade of declining expenditure. Caught by the twin pressures of an attack on public expenditure and the reducing size of the population of 18-year-olds over the period to 1996, the University Grants Committee (UGC) and National Advisory Board (NAB) are trying to organize and manage a cash-constrained and reducing system. Overlaying these problems is the increasing pressure for higher education to become more directly relevant to the needs of the economy and increase the share of expenditure going to the 'vocational' disciplines, and engineering and technology in particular. Yet not only has this switch to higher cost courses got to be made within a shrinking budget but the physical infrastructure has also to be brought up to date with increasingly expensive technology which itself can often become obsolete within a few years.

Last year the Organization for Economic Cooperation and Development (OECD) reported on the problem of an "ageing infrastructure" and obsolete equipment in many OECD countries¹. For twenty years, funding allocated for research has not matched the needs of the research community, and since the 1970s the cost of equipment has grown at a rate of 4 per cent above the inflation rate. In the United States it was estimated in 1982 that at least \$1,000 million was needed to bring university research equipment up to date. Many universities there now look to industry for help, and the tax regime makes donations tax deductible for companies. Although the majority of research and development funds for universities still come from government in one form or another, private industry has increased its spending on academic research and development by some 80 per cent since 1980 while government support has risen by only 34 per cent. It was estimated that US industry provided \$400 million to university research and development in 1984. Corporate motivation is clear: such investment aids graduate recruitment and gives access to teaching and research staff and facilities — all tax deductible. Several new university-based research centres have been established in recent years, such as that at Stanford, where 20 companies each paid \$750,000 towards new buildings, and then \$100,000 annually for three years to support educa-

tion, research and administration. New centres have also opened in Texas and North Carolina. Other ventures involve consortia of small companies, while others have matching state support.

In the United Kingdom the government sees this model as one way in which higher education can ease its funding problems. In 1985 it made available an additional £43 million under the 'Engineering and Technology programme' to provide an extra 5,000 undergraduate places in information technology and related subjects in the universities and polytechnics. A condition of part of that funding was that the recipient institutions should raise significant industrial support. In the event, assistance valued at over £24 million was forthcoming in the form of equipment donations, project work, student sponsorship, research and teaching by industrial staff. Industrial support is also given to many other courses and institutions, although it is impossible to assess the total contribution in the United Kingdom with any accuracy. In addition to this direct support for higher education, there is a growing programme of collaborative activities under national and international research programmes such as Alvey and Esprit, and those more directed at training, such as CASE awards and the Teaching Company Scheme administered by the UK Science and Engineering Research Council.

The majority of such links are negotiated at a personal level between individual academics and research or training staff in industry. As such, links have in the past developed on a very *ad hoc* basis. Now a number of major UK companies are seeking to develop corporate policies towards their support for higher education and are appointing 'campus coordinators' to draw together the company's activities with individual institutions in a more coherent way. They are still, however, a comparative rarity. While most collaboration is still on a one-off basis, with 'donations' often taking the form of payments in kind, there are some notable new jointly funded centres, such as the Leicester Biocentre which is funded by a consortium of companies to carry out fundamental research in biotechnology, and the recently announced Information Technology Institute organized by Cranfield Institute of Technology and supported by £4 million from over 20 high-tech companies. It

is to be a commercial research and training centre running both short courses and degree programmes. The most common form of industrial collaboration remains, however, the provision of project work for students.

There is a growing interest in collaboration in both higher education and industry, and a recent survey of information technology departments showed that 86 per cent wanted to increase the level of collaboration, the remainder being satisfied with their current activities. Ironically, a 'lack of academic staff and time' was acting as a constraint to increasing the level of collaboration in two out of three departments, yet such collaboration is seen as an essential way of boosting resources. There were also concerns by one in three departments that there was a lack of industrial interest or that the interest was too short term and piecemeal². Another problem was that out of date equipment reduced the attractiveness to industry, while many institutions still have over-bureaucratic procedures and uncertain guidelines by which academics can take on paid consultancy and outside work. In some institutions payments are supposed to be negotiated through the department with the individual receiving little direct cash reward, while in others staff are free to sell their time for personal gain. If industrial support is to grow and be seen as an integral part of an academic's timetable and reward structure, in part compensating for a low academic salary, then the institutions will need to be clearer and more positive towards the development and management of such links. And indeed some academics will have to start sharing their additional income with their full-time employers. It will take time for the links to develop as fully and as far as they have in the United States. The ultimate question for the United Kingdom is whether the industrial support will build on government funding or be seen as another reason for further reducing central expenditure on higher education and increasing the reliance on market forces. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Brighton BN1 9RF, UK.

1. *Science and Technology Policy Outlook* (OECD, Paris, 1985).
2. Connor, H. & Pearson, R. *Information Technology Manpower into the 1990s* (Institute of Manpower Studies, Brighton, 1986).